From DEPARTMENT OF MEDICAL BIOCHEMISTRY AND BIOPHYSICS
Karolinska Institutet, Stockholm, Sweden

STRUCTURAL AND FUNCTIONAL STUDIES OF PROTEINS IN CELL SIGNALING AND CANCER

Rebecca Dovega

Karolinska Institutet

Stockholm 2014

Published by Karolinska Institutet.
Original cover art courtesy of ShapeSciFx (http://www.shapescifx.com)
Printed by Universitetsservice US-AB

ISBN 978-91-7549-769-3

Structural and Functional Studies of Proteins in Cell Signaling and Cancer

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Rebecca Dovega

Principal Supervisor:
Professor Pär Nordlund
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Biophysics

Co-supervisors:
Dr. Christian Löw
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Biophysics

Dr. Rozbeh Jafari
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Biophysics

Opponent:
Professor Stefan Knight
Uppsala Universitet
Department of Cell and molecular biology
Division of Structure and molecular biology

Examination Board:
Dr. Christina Divne
Kungliga Tekniska Högskolan
Department of Biotechnology
Division of Biochemistry

Dr. Sonia Lain
Karolinska Institutet
Department of Microbiology, tumor and cell biology
Division of Cancer, tumor biology

Professor Sten Albert Ohlson
Nanyang Technological University
Department of School of Biotechnologies
Division of Structural Biology and Biochemistry

Till min älskade familj

ABSTRACT

Cancer is one of the most common causes of death in the western world. Cancer is an umbrella term for over 100 different diseases all caused by mutations in cells leading to uncontrolled cell division and metastasis. The control of the cell division is extremely important in order to maintain homeostasis and avoid development of cancer. The cells in the body have to communicate with each other to keep homeostasis. This can subsequently regulate the cell cycle that controls cell division. Many proteins are involved in the control mechanisms to pass the cell through the cycle where phosphorylation and dephosphorylation are important mechanism for regulation. The cell division is also controlled by tumor suppressors that sense, for example, DNA damage or cell stress, to ascertain reediness for cell division and to minimize the amount of mutations. Two important proteins in the cell signaling are Protein Phosphatase 2A that act as switch for the pocket proteins that control the cell cycle, and p53 which is labeled "the guardian of the genome", both are tumor suppressor.

Protein Phosphatase 2A (PP2A), is together with PP1 responsible for 90% of all dephosphorylation events in the cell. Apart from controlling the cell cycle PP2A is involved in many other pathways making it a very important protein in cell signaling. Understanding its structure and function is crucial for understanding how it works at the molecular level and how difference cancer mutations are affecting its mechanism of action. PP2A consists of three subunits, a scaffolding (A), a catalytic (C) and a regulatory (B) for substrate recognition. In this thesis the structure and biochemical function of one of the regulatory subunits, B"/PR70, is presented, a subunit who's structure was unknown before the start of the doctoral project. A high-resolution structure of the core revealed 8 EF-hands where two were binding calcium. A mapping of the A-B" interaction is also presented.

p53 is a common tumor suppressor that is mutated in 50% of all cancer tumors. It is involved in key decisions for cell cycle progression and apoptosis and more detailed understanding of this protein could shed light on the role of different p53 mutations in cancer. P53 is acting as a transcription factor and has a DNA binding core domain that binds to a response element on the target gene. In this thesis, a novel method to study biochemical events in cells and cell extract is applied for the first time on p53, the Cellular Thermal Shift Assay (CETSA). It is shown that p53-DNA can indeed be studied using this strategy and that binding profiles to four different oligonucleotides representing target genes with response elements, give distinct profiles for each mutation. This suggest that the CETSA strategy allows for more detailed functional studies of p53 in cells and that oligonucleotide profiling might constitute a novel mean to profile cancer patient cells for differential p53 activity.

LIST OF SCIENTIFIC PAPERS

I. **Rebecca Dovega**, Susan Tsutakawa, Esben M. Quistgaard, Madhanagopal Anandapadamanaban, Christian Löw, Pär Nordlund.

Structural and Biochemical Characterization of Human PR70 in Isolation and in Complex with the Scaffolding Subunit of Protein Phosphatase 2A. *PLoS One* (2014) Jul 9;9(7):e101846

II. Savitski MM, Reinhard FB, Franken H, Werner T, Savitski MF, Eberhard D, Martinez Molina D, Jafari R, **Dovega RB**, Klaeger S, Kuster B, Nordlund P, Bantscheff M, Drewes G
Proteomics. Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science*. 2014 Oct 3;346(6205):1255784

III. **Rebecca Dovega,** Rozbeh Jafari, Daniel Martinez Molina, Pär Nordlund. Biochemical Profiling of DNA binding to wild-type and mutant tumor suppressor p53 with CETSA. *Manuscript*

CONTENTS

LIST OF ABBREVIATIONS

AGF	Analytical gel filtration
CBLD	Calcineurin B-like domain
CD	Circular Dichroism
CDC6	Cell division control 6 protein
CDK	Cyclin-dependent kinase
CETSA	Cellular Thermal Shift Assay
DNA	Deoxyribonucleic acid
HEAT	Huntingtin-Elongation-A subunit TOR
His-tag	Histidine tag
IMAC	Immobilized Metal Affinity Chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
ITDRF	Isothermal dose response fingerprint
MDM2	Mouse double minute 2 homolog
MES	Minimal ensemble search
p53	Tumor protein p53
PP2A	Protein Phosphatase 2A
PSP	Protein Serine/Threonine Phosphatase
Rb	Retinoblastoma protein
RE	Response element
SAD	Single-wavelength anomalous dispersion
SAXS	Small Angle X-ray Scattering
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
TPP	Thermal Proteome profiling
TSA	Thermal Shift Assay

1 INTRODUCTION

About two billion years ago, eukaryotic cells appeared on earth. They have a nucleus containing all genetic information in chromosomes in contrast to prokaryotes that lack a nucleus. Eukaryotes can either be unicellular like amoebas and yeast, or multicellular like animals and humans. The human body consists of a huge number of cells, on average about one billion cells per gram of tissue. The cells exist in different forms in different tissues and organs. All these cells need to communicate with each other in order for the body to function properly and to maintain homeostasis.

Communication between cells is called cell signaling, and is a complex system to govern basic cellular activities and coordinate all cell actions. The communication occurs as primarily chemical signals, which includes neurotransmitters, hormones or growth factors. A cell releases a signal that reaches a target cell, the receptors on the target cell surface react to the signals and convert the external signals to internal signaling transduction pathways. An example of pathways are the MAPK signaling pathway, which is a phosphorylation cascade pathway that function to control cell proliferation and apoptosis, another example is Wnt signaling that is regulating development of cell proliferation [1].

Eukaryotic cells multiply by division. The cell cycle is divided in several stages or phases (figure 1) [2]. The cell is resting in G0 phase, it is quiescent i.e. out of cycle. When the cell is stimulated by growth factors or mitotic signals it moves into G1 phase, starts to grow and at a certain stage enters the S phase where DNA synthesis takes place, the whole genome is copied. Before the cell goes into S phase it passes a restriction point (R-point) controlled by the "pocket-protein" retinoblastoma (Rb) protein [3]. Rb is bound to a transcription protein called E2F, inhibiting transcription of genes regulating the S-phase entry, when Rb becomes phosphorylated it releases E2F and allows the genes to be transcribed. In gap 2 (G2) the DNA is controlled for completion and correct replication. Subsequently the cell starts to prepare for division, mitosis (M), where the cell is divided into two daughter cells with one set of chromosome pairs in each cell. The daughter cells have the same setup of DNA. The cell cycle then starts all over again with G1 or goes into G0 for rest. The cell cycle is regulated by cyclins and cyclin-dependent kinases (CDKs). There are a number of different cyclins involved in the cell regulation, the levels of the different cyclins are increased and decreased during the phases of the cell cycle. The cyclins are the regulatory unit of the CDKs that phosphorylates protein substrates to move the cell through the phases by activating DNA synthesis, and to form structural components needed for mitosis. For example CDK4/6 phosphorylates Rb for the release of E2F [4].

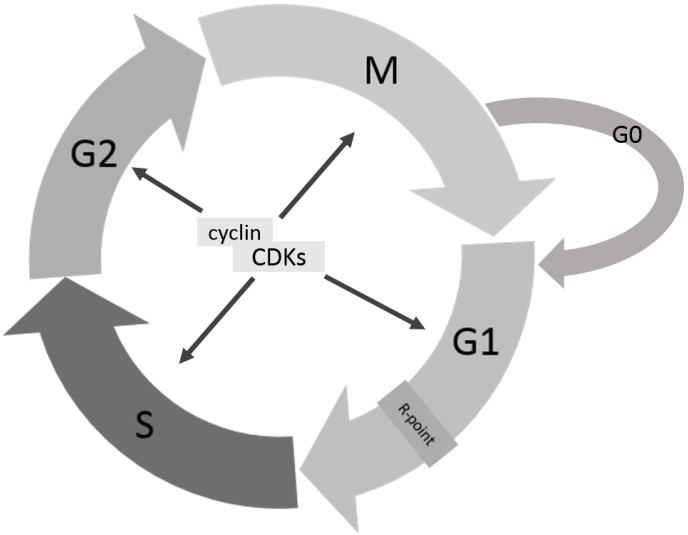

Figure 1. The cell cycle is divided in five phases. G0 the quiescence or rest phase. G1 were the cell starts to prepare for division, S phase with DNA replication, G2, and mitosis (M) when the cell is dividing. A restriction point (R-point) for control before entering the S phase. Cyclins and CDKs are regulators of the cell cycle.

When cells are no longer needed, they kill themselves with an intracellular death program known as programmed cell death or more commonly called apoptosis. Apoptosis also occur when the DNA get damaged by for example stress, UV-radiation or environmental chemicals. In order to protect the cell from mutations and diseases that could emerge from DNA damage, a signaling pathway starts in order to activate proteins that kill the cell. Apoptosis is therefore crucial for the multicellular organism to remove damaged cells. When DNA is damaged, p21, an inhibitor of CDK, is upregulated and stop the progression into S phase allowing the DNA to be repaired. If the DNA is severely damaged the tumor suppressor p53 activates pro-apoptotic proteins such as PUMA and BAX [5, 6] and the cell commits suicide instead of entering S phase.

When the cell fails to control the cell cycle and to initiate cell cycle arrest or apoptosis in response to cell damage, it leads to diseases, in many cases cancer. Cancer is the most common cause of death for people under 85 years-old in Sweden, and the number of cases are increasing [7]. Among women the most common form of cancer is breast cancer and for men prostate cancer. A third of all people will be diagnosed with cancer, but the survival rates are relatively high and improving.

Cancer is an umbrella term for over 100 different diseases involving abnormal cell growth with aggressive cells spreading to other parts of the body caused by metastasis [8]. Errors in the coordination of the cell cycle control lead to chromosomal alterations where genes are lost, rearranged or distributed unequally between the daughter cells,

leading to uncontrolled cell growth and tumorigenesis. Some tumors progress to cancer tumors by starting to spread their abnormal cells to other organs.

Improved cancer therapies are urgently needed. Many traditional ones are hitting a broad range of cells, both cancer and healthy cells. There are new therapies with drugs that directly target proteins which are more specifically activated in cancer cells, and therefore lead to less toxicity in normal cells. Although promising, drug resistance often develop rapidly to these and other therapies, and establishing concepts for drugs which can overcome resistance development, is a major challenge in cancer research. More detailed knowledge about how the cell works and how signaling and metabolic pathways are coordinated in cancer cells are needed to accomplish this. New concepts to screen for drug targets and drugs to find such therapies for specific types of cancers are also needed.

In this thesis two tumor suppressors, important for cell proliferation, are studied Protein Phosphatase 2A (PP2A) and p53. Structural and functional studies of PP2A shed light on its mechanism of regulation at the molecular level and studies of p53 provide a novel strategy to read out biochemical characteristics of this protein that might report on its activation state with implications for cancer therapy.

2 PROTEIN METHODOLOGIES

2.1 PROTEIN EXPRESSION AND PURIFICATION

To produce recombinant proteins, a host is needed to express the protein of interest in sufficient amount and in a well-folded and stable form. There are a range of different expression hosts available in both eukaryotic and prokaryotic cells. In the present work primarily *E. coli* has been used which is particularly useful for structural biology application, when it allows for rapid studies of many different expression constructs and conditions. One of the most commonly used *E. coli* systems is based on the BL21 (DE3) *E.coli* strain, with vectors containing T7 promoter. The gene coding for the protein of interest is inserted in a vector with T7 promoter, a purification tag such as e.g. 6xHis-tag, FLAG, or GST, and a protease cleavage site. A widely used expression system is the pET expression system, a variant of this is pNIC28-Bsa4 [9], which is used for ligation independent cloning (LIC)[10]. The empty vector contains a SacB fragment that is cleaved of when the gene of interest is cloned into the vector, and used for negative selection on agar plates. It also contains a gene for antibiotics resistance as positive selection. Colonies from the plates are inoculated to flasks with medium added with antibiotics, and grown to an OD_{600} typically in the range of 0.5-1 and then induced with IPTG. The induced cells are grown over night and then harvested and lysed for further purification of the protein.

It is critical to have a pure and well folded protein sample for successful crystallization of proteins for structural studies. In order to obtain that, the cell lysate with the protein of interest undergoes several purification steps; first an initial extraction from the *E.coli* lysate, often using a purification tag, and then purification using more elaborate methods such as Size-Exclusion Chromatography (SEC), Ion-exchange chromatography or hydrophobic interaction chromatography. For structural applications, the affinity captor of the protein is often combined with SEC when the later also is a good test of protein homogeneity.

Several purification tags can be used for protein purification, the His-tag is often used in purification from *E.coli* extracts as it provides good yields of proteins from high capacity resins [11] and does not affect the characteristics of the protein. The 6xHis-tag is attached with a protease cleavage site at the N- or C-terminal of the protein. The His-tag permits the protein to be purified by immobilized metal ion chromatography (IMAC) [12]. Metal ions like Ni^{2+} are immobilized on agarose resin packed in columns, and targets the histidines when the lysate is flowing through the column. To elute the His-tagged protein from the column a buffer containing high concentration of imidazole is added to the column (figure 2). The His-tag is usually cleaved off with TEV protease because its flexible nature might negatively affect crystal formation.

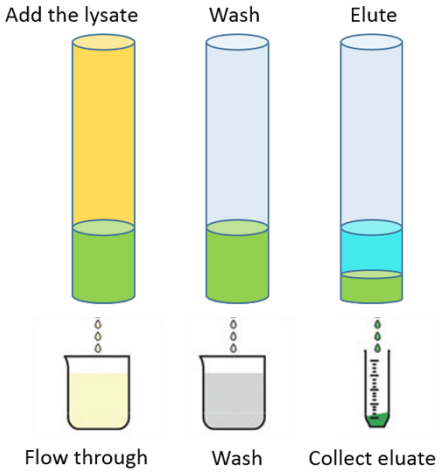

Figure 2. IMAC. The lysate is added to the column and the His-tagged protein will bind to Ni- beads. The flow through is then discarded while the column is washed with buffer to remove unbound E.coli proteins. Finally, the protein is eluted with high concentration of imidazole.

IMAC is usually followed by size exclusion chromatography to separate the protein of interest from contaminants and misfolded species. It is a method extensively used for protein purification as well as a control of protein folding/homogeneity. The protein solution is ran over a column packed with porous beads. Smaller protein will pass through the pores and larger proteins will pass by the beads in the column allowing for separation on size (figure 3).

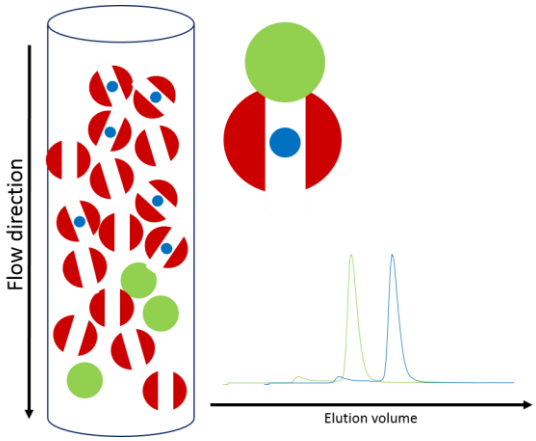

Figure 3. A cartoon illustrating the theory behind size exclusion chromatography. Small proteins (in blue) will pass through the beads (in red) in the column, the large proteins (in green) will pass on the outside of the beads and flow faster through the column and elute earlier than the small ones.

The area under the peak in the chromatogram is directly proportional to the amount of protein. A monodisperse peak is indicating good quality of the protein. If protein is not of sufficient quality after SEC, other purification approaches might be needed such as ion

exchange chromatography. However, a poor SEC profile might indicate that the protein, as produced in *E. coli,* is not in good shape and therefor, in practice, in this situation it is often worthwhile to test additional expression constructs of the protein and in other conditions.

2.2 ISOTHERMAL TITRATION CALORIMETRY

In mechanistic studies of proteins detailed measurements of interactions of the protein with physiological ligand such as a metabolite or a drug, or a small molecule are important. Isothermal titration calorimetry (ITC) is the most sophisticated method available to biochemists to study ligand interactions with proteins *in vitro*. The method measure the energy released or absorbed upon a physical interaction between two molecular components. The information acquired from an ITC experiment is enthalpy change, entropy change, binding affinity and stoichiometry. The ITC instrument has two cells, one reference cell supplied with constant power and a sample cell where power is supplied depending on the reaction that occurs (figure 4). When the ligand is titrated into the cell, the instrument detects the heat that is absorbed or released. It is done by measuring the changes in power to keep the sample cell and the reference cell in at same temperature, i.e. isothermal conditions.

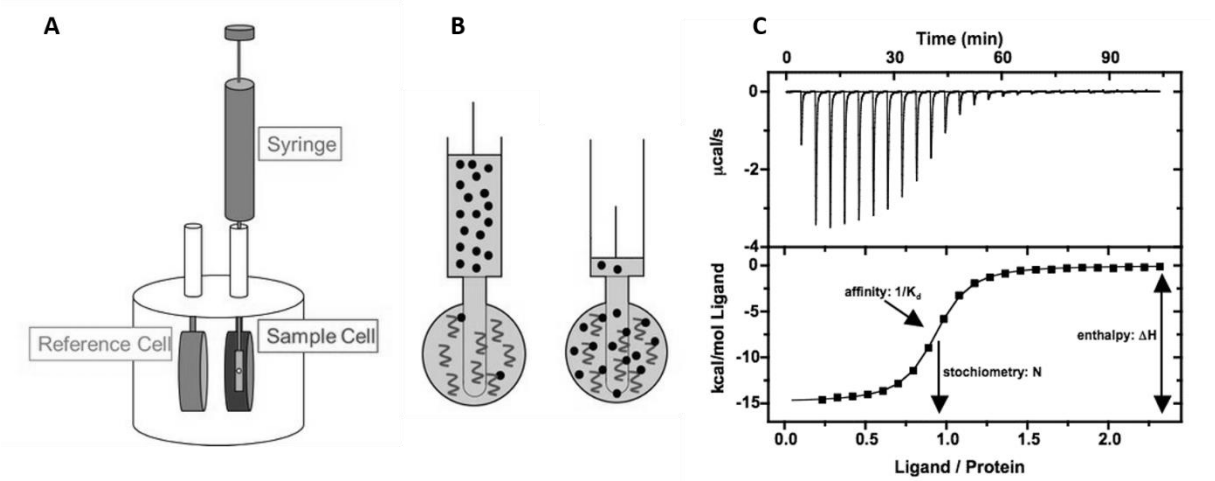

Figure 4. (A) Schematic picture of the instrument. (B)The ligand is stepwise titrated into the sample cell. (C)For every injection the heat needed to keep the sample cell isothermal is measured giving rise to a peak. The peaks are integrated to give a binding curve from which it is possible to calculate ⊿H, N and K_D.

The heat change is proportional to the amount of ligand bound and the binding enthalpy [13]. Therefore it is important that the concentrations of macromolecule in the sample cell and the ligand in the syringe are accurately measured for correct calculations. The first injection typically results in a large endothermal or exothermal signal because almost all ligand is bound to the macromolecule. For each injection the signal will be reduced as

the binding sites get more saturated. A typical exothermal ITC curve is shown in figure 4C. The reaction stoichiometry, N, and the dissociation constant is determined from the titration equivalence point.

There are several other methods available for binding studies of protein-ligand interactions. For example surface plasmon resonance (SPR) where one component usually a protein is immobilized on a sensor surface and a ligand is flown over the surface. Binding events are measured by monitoring change in resonance angle for the immobilized component [14]. Another method is weak affinity chromatography (WAC), where one component is immobilized on resin and another in the mobile phase. The stronger binding the longer the mobile phase will stay in the column [15]. WAC is good method for screening drug ligand libraries as it allows loading many ligands in the same time **[16].** The great advantage of ITC compared to other methods in measuring protein-ligand interactions is that it is possible to measure very small heat effects and therefore large binding constants up to 10^8-10^9 M^{-1}. In both SPR and WAC the protein needs to be immobilized which is not needed for ITC. That could interfere with the binding surface why ITC has an advantage over the other methods. Disadvantage of ITC is that relatively high concentration of protein is needed and that there is challenging to measure the binding of weak ligands.

2.3 ANALYTICAL GEL FILTRATION

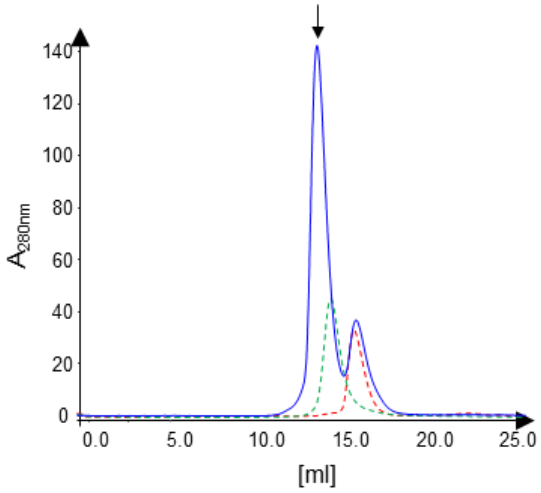

Figure 5. Analytical gel filtration with individual proteins shown in green and red, are eluting at different time points, when run together on the column they form a complex eluting with a shorter retention time and increased absorption height in blue.

Size exclusion chromatography can be used to study the interaction between two or more proteins. When using small volumes only for analysis the method is called analytical gel filtration (AGF). The individual proteins are first run one by one on the column to determine the elution volume of each individual protein. Then the proteins are mixed

with each other and run on the column. If the proteins interact with each other they will elute as a single peak at a shorter retention time than the individual proteins and with an increased absorption height (figure 5). In addition to studies of protein interactions AGF is also a valuable strategy to screen the homogeneity of membrane proteins after solubilization in different detergents [17].

2.4 LIMITED PROTEOLYSIS

Limited proteolysis is used to obtain a more stable truncated construct of a protein than an existing longer construct. It is often explored in structural biology to remove flexible regions to e.g. improve crystallization. A protease is used to cleave the peptide where flexible exposed regions will be cleaved off first. Therefore a more stable construct without flexible parts can generated. . The bands on SDS gels will show different stages where flexible parts are cleaved off. At the end there will be only one small band left, which is the most stable construct. The bands from the gel can then be analyzed with mass spectrometry to obtain the exact construct of the protein. In an alternative approach, a protease can be added directly to the crystallization trials to transiently generate highly crystallizable truncated proteins [18].

2.5 THERMAL SHIFT ASSAYS

Thermal shift assays (TSA) are used within many different methods like CD, fluorescence spectroscopy etc. in order to study stabilization of proteins upon ligand binding (figure 6).

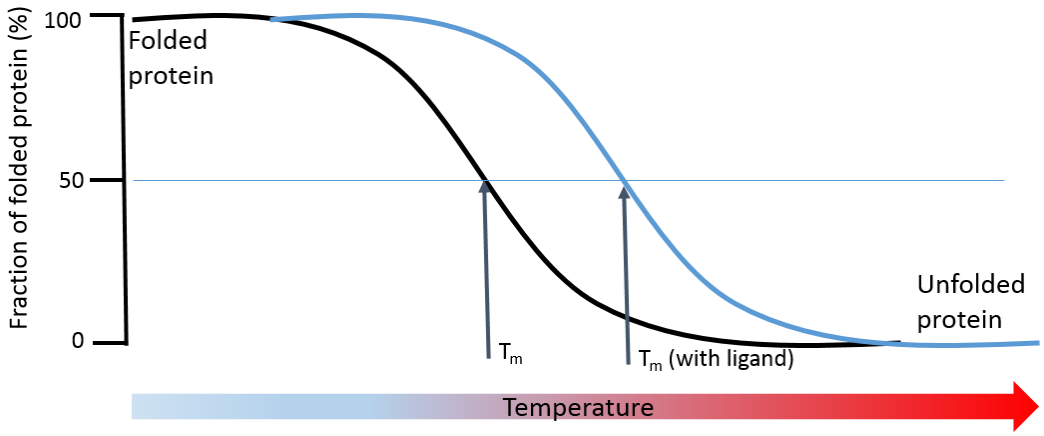

Figure 6. The protein denatures and unfolds when heated, the melting temperature is defined when half of the protein is unfolded. Upon adding ligand, the melting temperature is shifted (blue curve).

Circular dichroism (CD) spectroscopy can be used in many types of studies of chiral molecules but it is in the study of macromolecules like proteins it has been most useful. Circular dichroism (CD) is the difference in the absorption of left-handed circularly

polarized light (L-CPL) and right-handed circular polarized light (R-CPL). It occurs when a molecule contains one or more chiral chromophores (light-absorbing groups). CD is particularly good in studying the secondary structure of a protein and is therefore greatly influenced by the 3-dimension structure of the protein. Each structure has its own CD signature, see figure 7 [19]. This can be explored for structural study of a protein by predicting secondary structure content. The method is also useful for measuring conformational changes, including unfolding, due to change in temperature, pH, salt, ligands or denaturants. It is possible to use CD to measure the stability of a protein by determining the melting temperature, T_m, upon temperature increase. This is done by choosing a wavelength where the protein gives a strong or specific signal (like 222 nm for α helices) and measuring only at that wavelength while heating the sample. The melting temperature is the temperature where half of the protein is denatured.

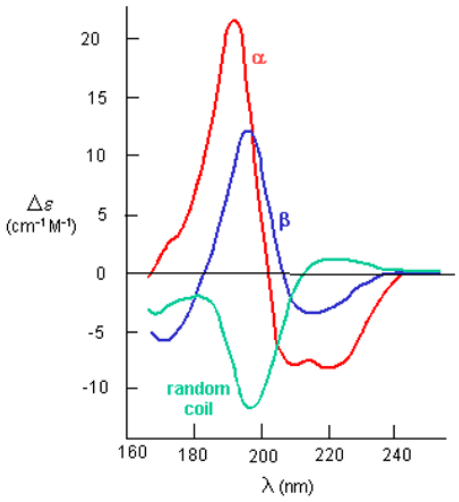

Figure 7. A CD spectrum of secondary structure, a helix, β sheet, and random coil.

Two frequently used assays in our lab is Differential Scanning Fluorometry (DSF) or thermofluor and differential static light scattering (DSLS), or stargazer [20]. In thermofluor the sample is mixed with a dye that binds to hydrophobic patches, when the protein is heated and unfolds the signal from the dye becomes stronger, the event is monitored with real time PCR. With DSLS the fact that proteins often precipitate after unfolding is explored and scattered light is used to measure the amount of aggregates formed. All these methods require recombinant pure protein, which sometimes could be problematic and expensive to produce.

Until recently it has not been possible to apply TSA in complex mixtures such as cells or cell extracts. Our group recently published a method based on the thermal shift concept in a cellular format, called Cellular thermal shift assay (CETSA) [21]. This assay allows,

for the first time, to directly measure the binding of a drug to its target protein inside cells. This can serve as a basis for understanding whether the cell is able to transport the compound across the cell membrane or if the compound needs modification/activation for protein binding. The CETSA approach also allows for studies of physiological interactions inside the cell and has the advantage that it can be directly applied to other complex samples such as tissues and excreted body fluids to measure biochemical events.

2.5.1 Experimental procedure of CETSA

In a CETSA experiment cells are cultured and then either intact cells or cell extracts are used for adding a ligand or as control experiments, without added ligand. The cells or cell extracts are then aliquoted into PCR-tubes and the different aliquots are heated in a temperature scale. If intact cells are used, they are lysed after the heating step by freeze thaw cycles. During the heating step the target protein denatures or precipitates at the higher temperatures and if the ligand is stabilizing the protein, it denatures and precipitates at a higher temperature than the protein without ligand. The precipitated proteins has to be separated from the remaining soluble folded protein. This is done by centrifugation or filtration. The supernatant is then loaded onto SDS-PAGE, followed by western blot. In the detection step an immunoassay such as a western blot could be used. The western blot signals for the target protein should be stronger at the lowest temperatures and then start to decrease as the temperature increases. Plotting the intensities of the western blot signals against temperature will give a melt curve. By normalizing the intensities it will be possible to read out the T_m, which is the temperature where half of the protein is denatured.

The CETSA method is also applicable for dose response experiments, which is often used in drug discovery to measure how efficient a drug is in different settings. It is the same procedure as mentioned above for obtaining CETSA melting curves, but instead of having a temperature scale the compound is added to intact cells or cell extracts at different concentrations. All the samples are then heated at the same temperature, which derived from the melt curve by taking a temperature where a strong difference is between the sample with and without ligand. After heating, the cell extracts or cells are processed as for the melt curve experiments. The band intensities are here normalized and plotted against the ligand concentration to give an isothermal dose response fingerprint ($ITDRF_{CETSA}$) curve. The concentration at half-saturation is not the exact affinity, but typically shifted slightly [22], but it is characteristic fingerprint which can be used to e.g. compare different cell types, compounds etc. [21, 23]. For a more high-throughput procedure, it is possible to use other detection methods like AlphaScreen® or mass spectrometry [23].

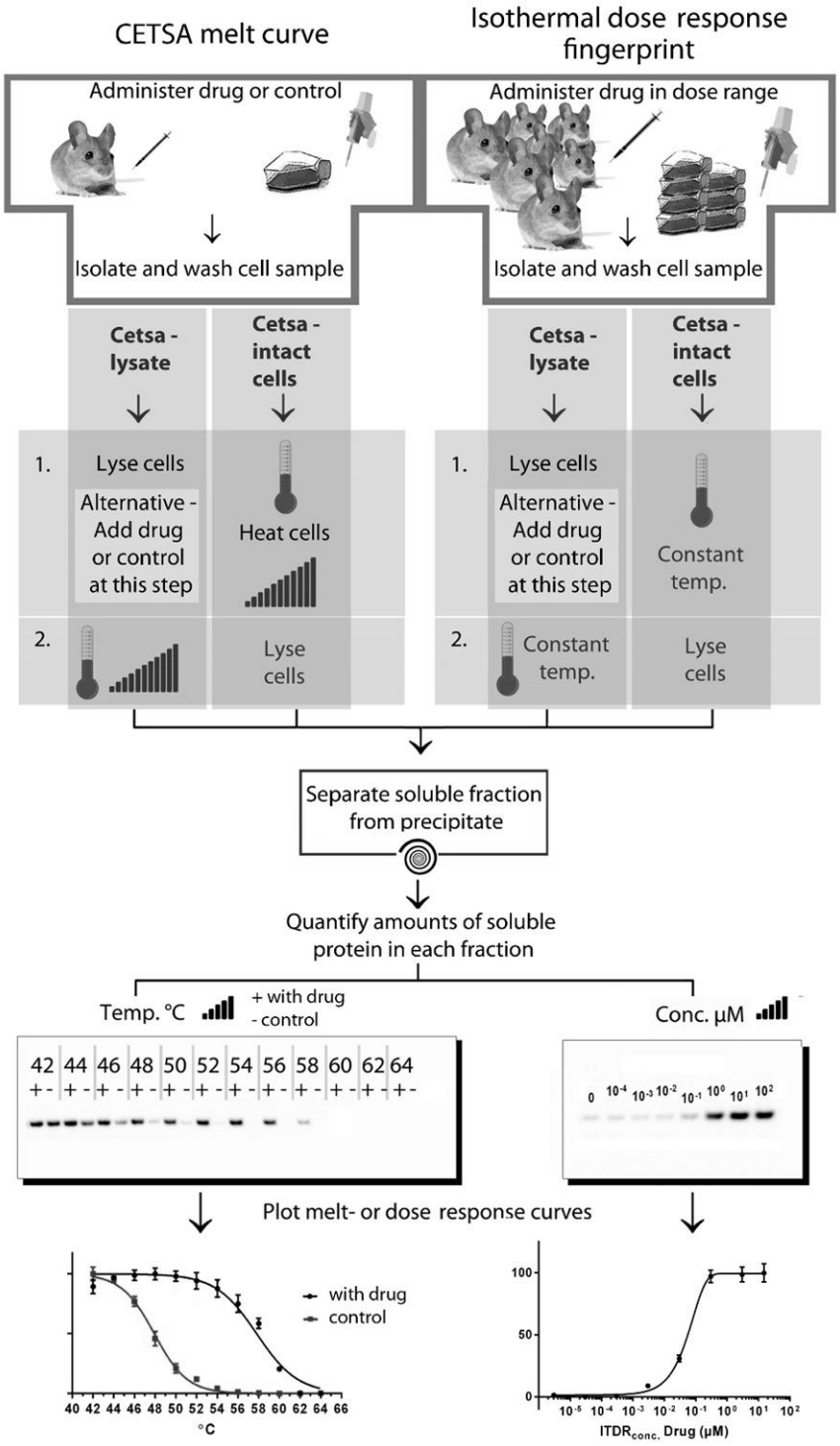

Figure 8. Schematic illustration over CETSA melt curve and ITDRF$_{CETSA}$ procedure.[21]

3 STRUCTURAL BIOLOGY

Structural biology is the study of the molecular structure of biological macromolecules, especially proteins and nucleic acids, their three-dimensional molecular structure, how they acquire the structures, and how the structure affects their mechanism of action. This subject is of great interest to biologists because macromolecules carry out most of the functions in cells, and by folding into specific three-dimensional shapes, they are able to perform these functions.

There are several methods to study protein structures, the most popular is X-ray crystallography, but nuclear magnetic resonance (NMR) spectroscopy and electron microscopy (EM), either of 2-dimensional crystals or single particle reconstitutions, also play important, but partly complementary, roles. These mentioned methods can yield high-resolution structures of proteins but a complement to these methods for lower resolution studies in solution is small angle X-ray spectroscopy (SAXS).

3.1 X-RAY CRYSTALLOGRAPHY

Already in 1934, J.D. Bernal and Dorothy Hodgkin showed that proteins packed in crystals diffracted X-rays in a complex but regular patterns. These patterns contained the essential information needed to determine a protein's structure. This could however not be accomplished until Max Perutz and colleagues compared different patterns containing different heavy-metal atoms. The first protein structure was solved by Kendrew et al in 1958 and was myoglobin from sperm whale [24].

X-ray diffraction is a common physical phenomenon and occurs when a wave (of any nature) encounters an obstacle, which can be any material object. This results in bending of the wave around that object, also called scattering of waves. X-rays are electromagnetic waves with a wavelength of around 1 Å (i.e. 10^{-10} m, 0.1 nm). Diffraction from atoms in a crystal are following Bragg's law, the pattern produced gives information about the distance between the crystal planes. To collect a typical diffraction pattern from proteins, the beam of the x-rays should be monochromatic (only one wavelength), high intensity, and highly focused. While rotating the crystal in the beam diffraction images are collected. An example is shown in figure 9. Each diffraction spot corresponds to a point in the reciprocal lattice and represents a wave with an amplitude (F) and a relative phase (α) used in the process to generate the experimental electron density. The structure factors for the various points on the reciprocal lattice correspond to the Fourier transform of the electron density distribution within the unit cell of the crystal. The Fourier transform is reversible, which means one can go from electron density to reciprocal phase and back again. The problem with going from reciprocal space to electron density is the missing phase information required for the generation of the electron density, this is called the phase problem.

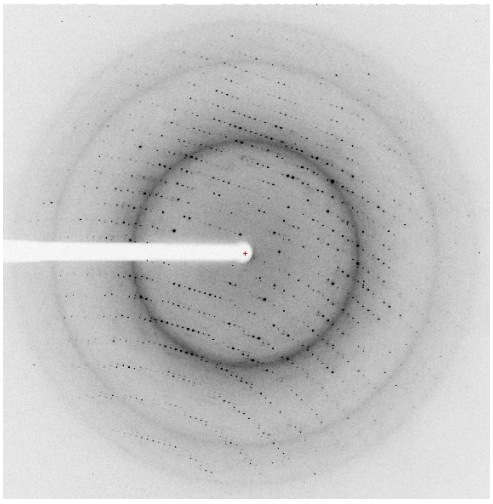

Figure 9. Diffraction pattern from a protein crystal. Each diffraction spots corresponds to point in the reciprocal lattice.

The phase problem can be solved with several methods.

The most straight forward is **molecular replacement**, which can be used in cases where there is an homologous structure, preferably with at least 35% sequence identity [25]. The known structure model is used as an initial model for phase generation after the rotation and translation problem has been solved, which allow the positioning of the molecule in the crystal lattice.

Isomorphous replacement and **Anomalous dispersion** are both based on the present of a heavy atom in the protein that is scattering the X-rays in a different way than carbons, oxygen and nitrogen. For isomorphous replacement a diffraction pattern of the native protein crystal and a derivative is needed, for anomalous dispersion a diffraction pattern of the derivative is enough. For isomorphous replacement it is important that the heavy atom does not change the structure of the protein i.e. the unit cell has to be the same. For anomalous dispersion methods the measurements of the diffracted intensities have to be very accurate, to allow small differences between, so called, Friedel pairs to be derived which can be used to calculate the initial phase information.

3.2 SAXS

Small-angle X-ray scattering (SAXS) is used to study biological macromolecules in solution. SAXS provides information about the overall structure and structural transitions of native biological macromolecules in solutions at low resolution (1–2 nm). The experiment is more straightforward as compared to crystallography; there is no need for crystallization, isotopic labelling, or heavy atoms. The measurements are performed close

to the primary beam, hence the name "small-angle". The sample solutions are exposed to monochromatic X-rays and the sample is scattering the waves, the solvent is measured separately to subtract the solvent scattering from the sample. The net intensity is used to extract the structural information, it is proportional to the scattering from a single particle averaged over all orientations. Using the precise scattering patterns recorded at synchrotrons, SAXS can reveal the shapes and weight of protein molecules and nanoparticles e.g., radius of gyration Rg, maximum diameter D_{max}, volume of the hydrated particle V_p, and the molecular mass of the particle. Moreover, the low resolution macromolecular shape can be obtained *ab initio* (i.e., without information from other methods), and molecular models can be constructed using high resolution or homology models of individual domains or subunits.[26]. However, care has to be taken with generated structures when it sometimes is hard to distinguish that right structure model from false structure models.

To process the data several programs are available, to create atomic models and fit the experimental scattering data. For example ATSAS is based on molecular dynamics, where structured regions can be maintained as rigid bodies, and flexible loops and termini are allowed to move, bond lengths are kept constant. Conformational sampling is followed by structure validation and a minimal ensemble search (MES), where the program will identify ensembles that together will be similar to the scattering curve.

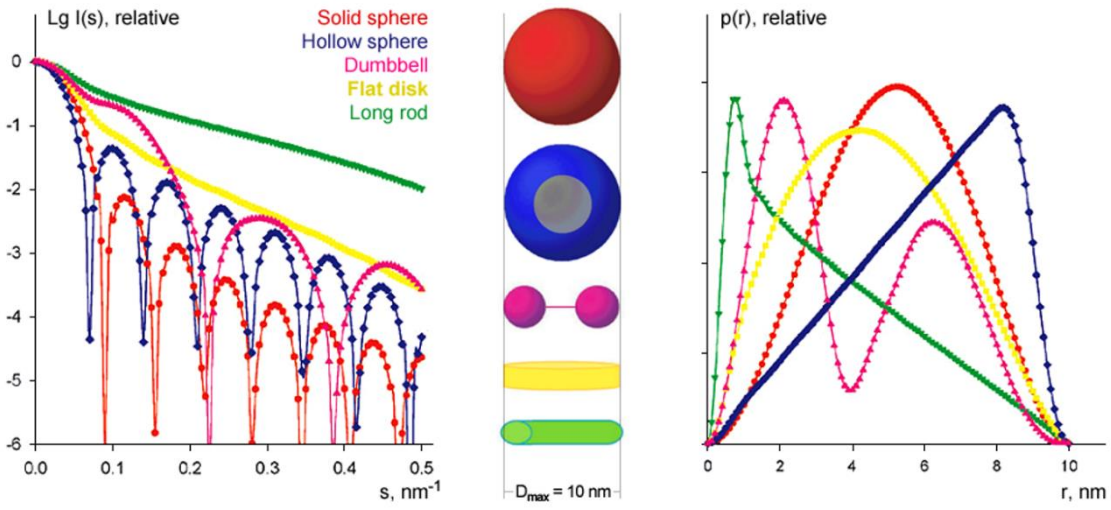

Figure 10. Scattering intensities and distance distribution functions of geometrical bodies. [27]

4 PROTEIN PHOSPHATASE 2A

4.1 SERINE/THREONINE DE-/PHOSPHORYLATION

Reversible phosphorylation and dephosphorylation events are essential for the regulation of a multitude of cellular functions [28, 29]. Protein phosphorylation was first introduced by Edmond Fischer and Edwin Krebs [30-32] when they could demonstrate the process of a protein transferring a phosphate group from ATP to phosphorylase *b* resulting in phosphorylase *a*. The converting protein was named phosphorylase kinase. Another protein performed the reverse conversion, dephosphorylation, that protein was named phosphorylase phosphatase[33].

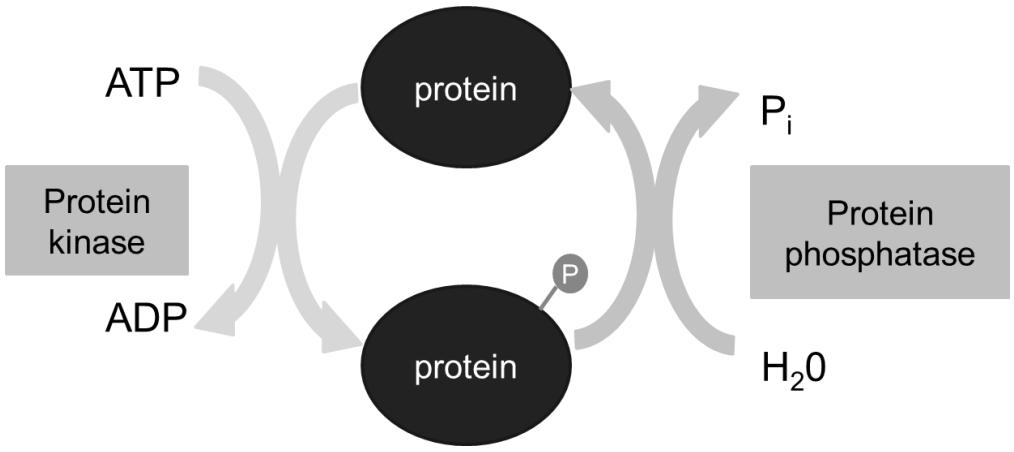

Figure 11. Kinases and phosphatases are working together by activating and deactivating proteins via phosphorylation and dephosphorylation respectively.

The protein family that is responsible for transferring a phosphate from ATP to a protein is named protein kinases, and the family responsible for the opposite is called protein phosphatases (figure 11). In eukaryotic cells de-/phosphorylation mainly occurs on hydroxyls of serine, threonine and tyrosine residues of substrate proteins. The human genome contains only ≈30 protein Ser/Thr phosphatases (PSPs) in comparison to 428 Ser/Thr kinases (PSKs). The reason for this difference is that holoenzymes of PSPs are generate by combinations of regulatory subunits interacting with a limited number of catalytic subunit, discussed further below.[8]. The PSPs are divided in three major groups: phosphoprotein phosphatases (PPPs), metal-dependent protein phosphatases (PPMs), and the aspartate-based phosphatases represented by FCP/SCP (TFIIF-associating component of RNA polymerase II CTD phosphatase/small CTD phosphatase)[8]. For several members of the PPP family, the catalytic subunit associates with a great variety of regulatory subunits. Representative members of the PPP family include protein phosphatase 1 (PP1), PP2A, PP2B (commonly known as calcineurin), PP4, PP5, PP6, and PP7.

4.2 CELLULAR ROLL OF PP2A

The cell cycle is regulated in a complex system initiated by many different pathways such as e.g. MAPK and Wnt signaling. Protein phosphatase 2A (PP2A) participates in various pathways controlling metabolism, DNA replication, transcription, RNA splicing, translation, cell cycle progression, morphogenesis, development and transformation [34-37]. Protein phosphatase 2A is regulating other proteins either negatively or positively by dephosphorylation. As an example, PP2A regulates so called pocket proteins, which negatively regulate the cell cycle by targeting transcription factor E2F. E2F then modulate the expression of genes involved in cell cycle progression. Pocket proteins consist of p107, p130 and pRb [38]. They are active in their hypophosphorylated state, and inhibited when they are phosphorylated by cyclin/CDK complexes [39]. Active pocket proteins bind to E2F inhibiting DNA synthesis (Figure 12). When pRB/p107 is phosphorylated, E2F gets released and expresses genes allowing entry to S phase [40]. Different cell stress signals activate PP2A to dephosphorylate pocket proteins leading to cell cycle arrest or apoptosis. For example oxidative stress induces an increase of Ca^{2+} intracellular levels activating PP2A, which leads to dephosphorylation of pRb [38].

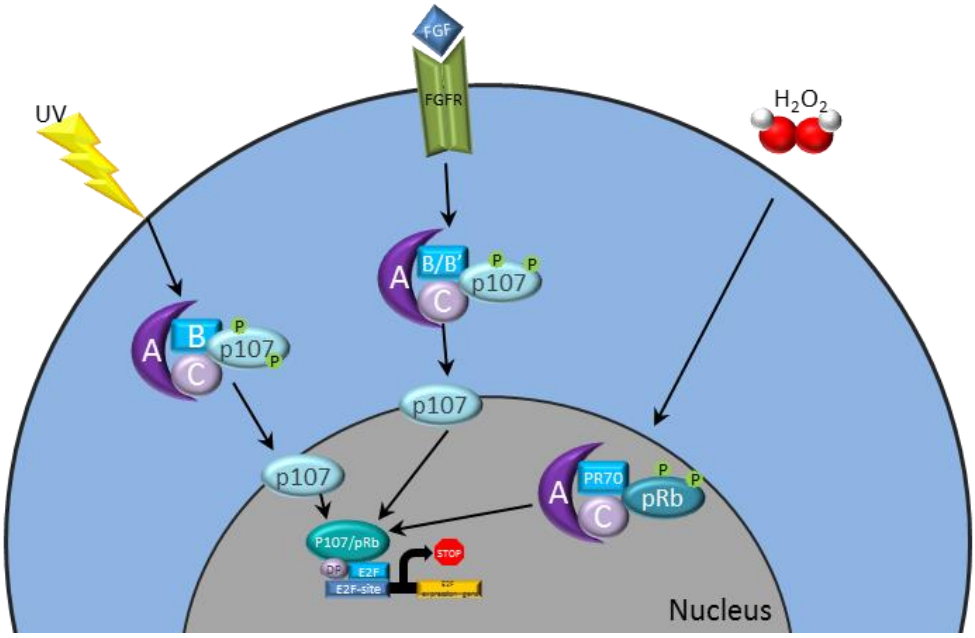

Figure 12. PP2A have several roles in the cell, one of them is to dephosphorylate pocket proteins (p107 and pRb) when the cell is stressed from UV-radiation, ligands or hydrogen peroxide. The pocket proteins are stopping the cell to enter S phase by binding to E2F.

Another substrate of PP2A is Cell Division Control 6 (Cdc6) protein, which contains three phosphorylation sites. Cdc6 is the gate keeper for the G_1/S transition and is required for DNA replication [41]. When Cdc6 is dephosphorylated by PP2A PR70 it allows the cell to proceed from G1 to S phase initiating DNA replication. Interestingly

high levels of Ca^{2+} increases association of Cdc6 with the catalytic and scaffolding subunits of PP2A but does not affect the PR70/Cdc6 interaction [42].

Protein phosphatase 2A has been suggested to be a tumor suppressor [43] consistent with that okadaic acid, a potent tumor promoter, is strongly inhibiting PP2A. This was verified after Wang *et al.* found that the gene coding for the β isoform of A subunit was altered in 15 % of primary lung tumors, and 15 % of primary colon tumors [44]. Several mutations in both α and β isoforms have been found in diversely different cancer types [45]. In Aαisoform, four point mutations namely E64D, E64G, E393Q and R418W are found in cancer. In Aβ isoform, eight point mutations have been found in cancer patient, they are G8R, P65S, G90D, L101P, K343E, V448A, D504G, and V545A [46].

4.3 THE HOLOENZYME OF PP2A

Protein phosphatase 2A (PP2A) is a major Serine/Threonine phosphatase that together with Protein phosphatase 1 accounts for more than 90 % of all dephosphorylation processes in any given tissue or eukaryotic cell. Each holoenzyme consists of a common core formed by the scaffolding (A) and the catalytic (C) subunit (two isoforms each) and associates with a variable regulatory B-subunit to form a heterotrimeric complex (figure 13).

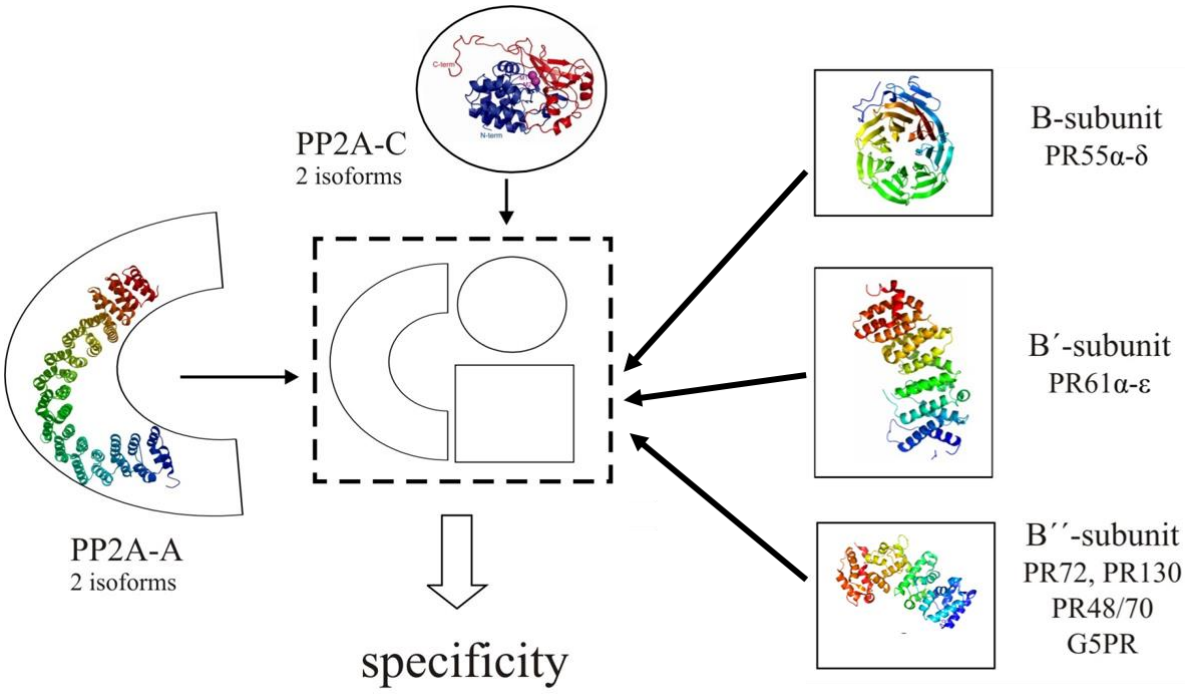

Figure 13. PP2A holoenzyme consists of three subunits, a scaffolding (PP2A-A), a catalytic (PP2A-C) and several regulatory subunits (PP2A-B) divided in three families. The B-subunits recognizes the substrate giving PP2A substrate specificity.

4.3.1 The scaffolding subunit

PP2A-A, the scaffolding subunit, consists of 15 tandem HEAT (Huntingtin-Elongation-A subunit TOR) repeats that formed an overall conformation similar of a hook [47]. Each HEAT repeat is formed by two antiparallel α-helices, with inter-helical loops that are highly conserved through evolution. Structural studies have shown that HEAT repeats 11-15 are involved in binding the catalytic subunit [48] while the ridges in HEAT repeats 1-8 are involved in binding the regulatory subunit [49, 50]. When binding to catalytic subunit, HEAT repeats 13-15 are moving 20-30 Å resulting in a drastic conformational change between repeat 12 and 13, this gives the scaffolding subunit a more horse-shoe like shape [48]. When forming the holoenzyme the N-terminal repeats of the scaffolding subunit are moving 50-60 Å, the changes and the flexibility in the scaffolding subunit may facilitate the phosphatase activity by moving the bound substrate protein on regulatory subunit, closer to the catalytic subunit. The scaffolding subunit exists in two different isoforms; α and β, they have 85% sequence identity but binds with remarkably different affinities to the regulatory and catalytic subunits [46]. β isoform is expressed in much lower levels than α isoform in the cell [51]. The β isoform also has 12 extra amino acids in the N-terminal that does not exist in the α isoform, but the exact role of these residues is still uncertain.

4.3.2 The Catalytic subunit

PP2A-C, the catalytic subunit, exists in two different isoforms; α and β. Other phosphoprotein phosphatases (PPPs) share together with PP2A a conserved fold as well as the structure essential parts of the active site [8]. However catalytic subunits of other PPPs do not bind to the scaffolding subunit of PP2A. Structural analysis of the interface between the catalytic and scaffolding subunits shows that the amino acids in the catalytic subunit that are important for interaction are not the same in PP2A and other PPPs. (ref) The catalytic subunit of PP2A has a highly conserved carboxyl-terminal tail, with a TPDYFL motif. This is important for the holoenzyme assembly and in the selection of proper regulatory subunits. Post-translational modifications of the tail, including methylation of leucine and phosphorylation of tyrosine, are also adding further control of B regulatory subunit selection [52].

4.3.3 The Regulatory subunit

There are four families of regulatory B-subunits with no homology between them. They have very different expression levels in different cell types and tissues. Examples of B-subunits are: B/B55/PR55, B'/B56/PR61, B"/PR72 and B"'/Striatin/PR93 [53, 54]. Within the holoenzyme, the regulatory B-subunit controls the function of PP2A by mediating substrate specificity and modulating the catalytic activity. Crystal structures of PP2A holoenzymes with members of the regulatory B, B' and recently B" subunit

families [41, 49, 50, 55, 56] have shed light on the overall enzyme architecture and gave initial ideas on how regulatory subunits might affect PP2A activity.

The regulatory subunits from different families are not related in their structures and therefore show no sequence homology. B/B55/PR55 forms a seven-bladed β-propeller; each blade is a WD40 repeat [57] (a repeat of 40 residues ending with a tryptophan and an aspartate residue) with four β-strands. At the top face of the propeller, there is a highly acidic groove; this groove is predicted to host the substrate-binding site. A key function of PP2A-B55 is to dephosphorylate the hyperphosphorylated Tau protein, preventing Tau to polymerize into neurofibrillary tangles in the brain, which is a process implicated in Alzheimer's disease [58-60]. The bottom face of the β-propeller binds to the ridges of HEAT-repeat 3-7 in the scaffolding subunit, and an extended β-hairpin arm reaches down to interact with HEAT-repeat 1 and 2 (figure 13) [50] B/B55/PR55 makes few interactions with the catalytic subunit; the FYDL-motif was disordered in the crystal structure, indicating that methylation of this motif is not essential for holoenzyme assembly in this family.

B'/B56/PR61 has a similar fold as the scaffolding subunit, consisting of eighteen α helices where sixteen of them are forming eight HEAT-like repeats. The protein has an elongated structure with an apparent curvature. The convex side contains a number of conserved residues forming interactions with HEAT-Repeat 2-8 of the scaffolding subunit [49]. The potential substrate-binding site is acidic and positioned on the concave surface of the regulatory subunit B', close to the active site of the catalytic subunit [8].

B"/PR70/PR72 had until recently an unknown structure. The only predicted structural elements were two EF-hands involved in Ca^{2+}binding. Structure and interaction with the scaffolding subunit will be described further down.

4.4 STABILITY STUDIES OF SCAFFOLDING SUBUNIT ALPHA AND BETA ISOFORM

As mentioned above, many cancers show mutations of the scaffolding subunit of PP2A. It has also been demonstrated that these mutations affect the formation of the holoenzyme complex [61]. The stability of a protein can give information about the effect of the mutations. To study the thermal stability of both the α and β isoform wild-type and mutant proteins, we performed a study where the PP2A scaffolding subunits were heated from 20 to 70 °C and measured with CD at a wavelength typical for α helices (figure 14). These experiments revealed that the β isoform is more than 10 °C less stable than the α isoform in spite of their high sequence similarity. We also showed that all cancer related mutations in β isoform are destabilizing the scaffolding subunit, except L101P and D504G that give a small stabilization of 1 °C. The later mutations are located at the interacting site of B and the C subunits, hence affecting the binding and not the

intrinsic stability. The most destabilizing mutant was K343E, which is situated between HEAT repeat 10 and 11 forming the hinge region between the two parts binding the B and C subunit respectively. For the α isoform none of the mutants was found to destabilize the scaffolding subunit.

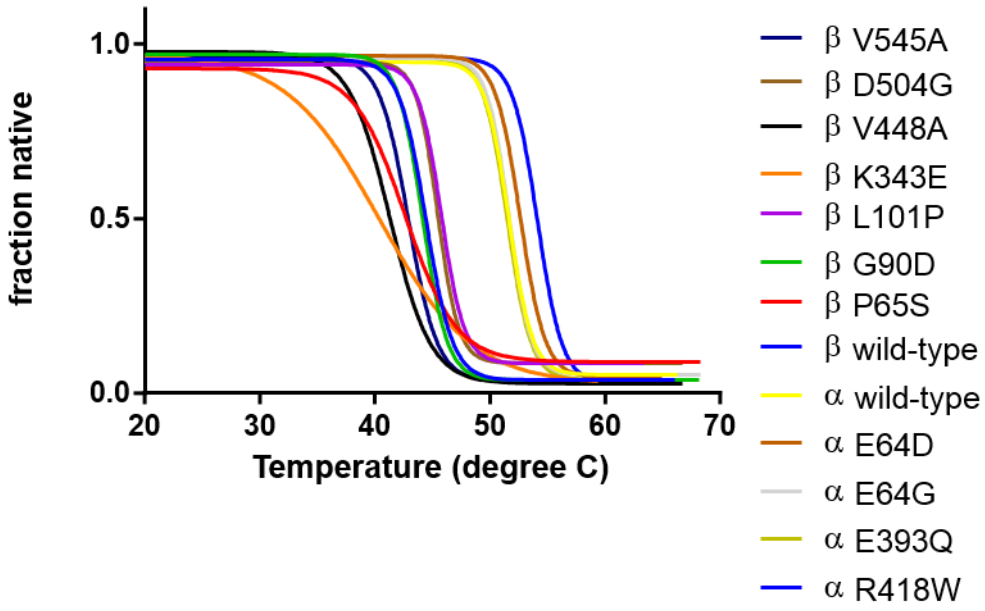

Figure 14. Melt curves of scaffolding subunit a and β isoform of wild-type and mutants. a isoform is significally more stable than β isoform.

4.5 STRUCTURAL AND BIOCHEMICAL CHARACTERIZATION OF HUMAN PR70 IN ISOLATION AND IN COMPLEX WITH THE SCAFFOLDING SUBUNIT OF PROTEIN PHOSPHATASE 2A (PAPER I)

4.5.1 Construct design

No structure was known for the heterotrimeric complex of A/C/B" until recently [41]. The missing piece in the puzzle of PP2A had to be found. The catalytic subunit of PP2A was impossible to express in *E.coli* and in our hands, and only low levels could be produced in insect systems. Luckily, the A/B"-complex was possible to produce and assemble without the C subunit, this is not possible for any other of the other studied B-subunits. Several crystallization attempts of the full-length constructs of A and B"/PR70 in complex failed. Subunit A was able to crystallize by itself but together with B"/PR70, there were no crystals. Full length B"/PR70 didn't crystallize by itself either. Instead we designed constructs that still had the predicted interacting regions but not the flexible regions of the A subunit, .It was known that HEAT-repeats in the N-terminus are involved in the interaction with the regulatory subunit, so by first deleting repeat 11-15

and subsequently one HEAT-repeat at a time from the carboxyl terminus, 10 new constructs were designed, where four allowed soluble protein to be produced. These were for HEAT repeats 1-7, 1-8, 1-9 and 1-10. These protein constructs were analyzed on analytical gel filtration (AGF) with full length B"/PR70 to see which constructs were still able to form the complex. All except repeat 1-7 retained the complex forming abilities. The new constructs of Aα formed a complex with full length B" did unfortunately not crystallize either.

A more stable B"/PR70 construct without flexible regions could potentially help, to obtain this limited proteolysis was performed. Full length B"/PR70 was mixed with chymotrypsin, and aliquots were taken at time points between 0 and 360 minute and studied by SDS-PAGE. Three different bands were identified on the gel. The bands were cut out from the gel and analyzed with mass spectrometry. Three new constructs were designed with construct boundary of 25-483, 130-575 and 130-483. Preceding these experiments, a construct of residues 105-575 already existed. The new constructs were run on AGF together with full length and HEAT 1-8 of A subunit, revealing that constructs 25-575 and 105-575 was making a nice complex. Unfortunately none of the complexes crystallized. Therefore, crystal screens for the different B"/PR70 constructs alone was also set up.

4.5.2 Crystallization and structure determination of PP2A PR70

The core domain with residues 130-483 was the only construct yielding well diffracting crystals. No molecular model was available for phasing at that time, so the protein was cultured in minimal media with seleno-methionine to be used for anomalous dispersion experiments. Data were collected at the selenium edge at Diamond Light Source beamline I04, a tunable beam line with Pilatus detector. The phasing was done with single-wavelength anomalous dispersion (SAD), where differences in diffracting intensities of anomalous pairs, were used to identify selenium sites, and to calculate initial phase information. In total 10 selenium sites were found and the initial structure was built and subsequently refined using Phenix, with manual rebuilding in Coot in each cycle of the refinement.

4.5.3 Overall structure of PR70

The crystal structure of PR70 revealed an elongated protein with eight EF hands. The protein can be divided into two domains connected with a 12 residue long linker (figure 15). Except for two EF-hands there was no previous structure predicted for PR70. A search on the amino acid level did not return any structure homologues. However a DALI search on the structure, returned a hit for the regulatory subunit of calcineurin (calcineurin B), each of the domains of PR70 have a similar fold as calcineurin B and was

therefore named Calcineurin B-like domains (CBLD). A CBLD consists of four EF hands, CBLD1: EF hands 1-4 and CBLD2: EF hands 5-8

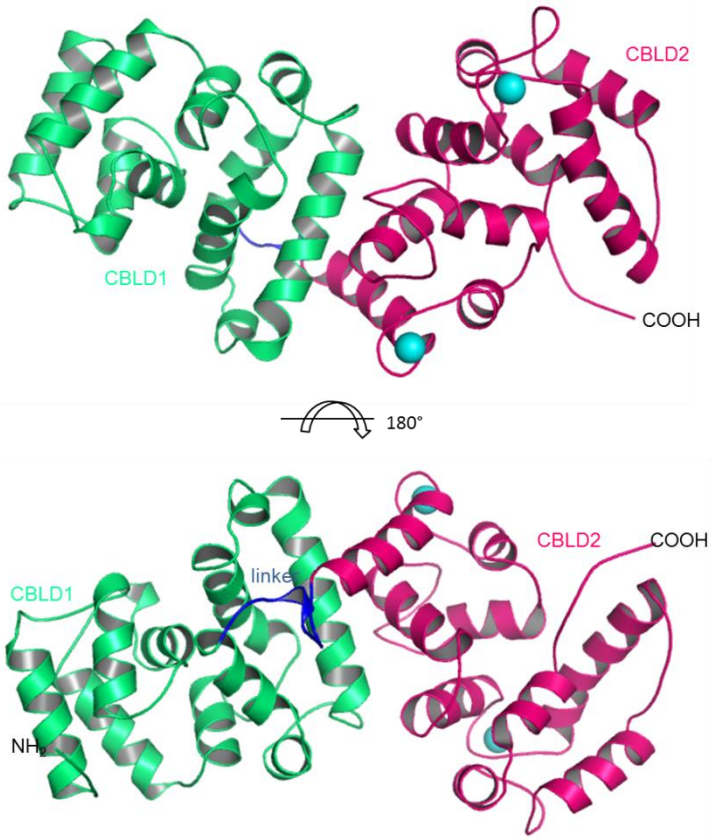

Figure 15. The overall crystal structure of PR70 core fragment. CBLD1 is shown in mint green and CBLD2 in magenta, the linker is shown in blue.

The EF-hand is a helix-loop-helix motif (E-helix, EF-loop, F-helix), which often has the capacity to bind calcium. The EF-loop is twelve residues long and has a consensus sequence with aspartate residues in positions 1, 3, 5, glycines in positions 4, 6, and a glutamate in position 12 (table 1) [62, 63]. EF-hands tend to come in pairs forming a discrete domain; many come in two, four, six or eight EF-hands, like is seen in the PR70 core domain. The two EF-hands in a pair interact with each other via small β-strains in the EF-loop (figure 16A). In agreement with previous observations, only two of the EF hands of PR70 are binding calcium [38, 41, 42]. These EF-hands correspond to EF5 and EF7 in the structure (figure 16B). These were previously referred to as EF1 and EF2 respectively in the literature when it was not known that the structure contained more than two EF-hands.). The major differences between EF5 and EF7 is the aspartate and the glutamate respectively in position 12, and the lack of glycines in position 4 and 6 in EF5 (table 1). The substitution of the canonical glutamate for an aspartate in position 12 usually leads to a more compact EF-hand as the shorter side chain has to reach further to interact with the Ca^{2+} ion [63-65]. However, in EF5, the aspartate does not interact

directly with the Ca^{2+} ion, but forms a single indirect interaction mediated by a water molecule. The structure thus agrees well with the suggestions that EF5 binds calcium with lower affinity than EF7 [41, 42]. All other EF-hands of PR70 are consistently missing the aspartates in position 1, 3 and 5 and therefore it is not surprising that none of these EF-hands are binding calcium.

Table 1. Sequence of EF loops 5 and 7. EF loop consensus sequence is shown. Residues conforming to the consensus are highlighted in grey.

Position	1	2	3	4	5	6	7	8	9	10	11	12
Consensus	D		D	G	D	G		I				E
EF5	D	T	D	H	D	L	L	I	D	A	E	D
EF7	D	L	D	G	D	G	A	L	S	M	F	E

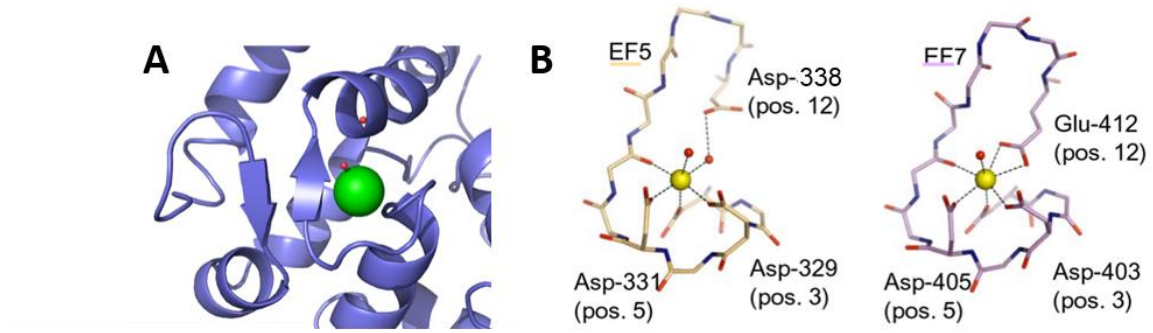

Figure 16. A) Two adjacent EF-loops are interacting with each other via a small β strain. B) Structure of EF loops 5 (pale orange) and 7 (violet). Water molecules interacting with calcium are red. Side chains are omitted for residues not directly involved in calcium binding (except for Asp-338).

4.5.4 Binding of Ca^{2+} to PR70

The crystal structure of PP2A PR70 revealed two Ca^{2+} ions bound to two EF-hands. This had previously been described as a unique feature for the B"-regulatory subunit family [66]. To further investigate the role of calcium we used a thermal shift assay to study binding of Ca^{2+} to PR70. Thermal shift assays were measured using circular dichroism of PR70 (residues 105-575) with different concentrations of calcium. When 50 μM calcium was added, there was a clear shift of T_m, from 29°C to 41°C (figure 17). Overloading PR70 with calcium (500 μM) a second shift could be seen. Plotting the T_m against calcium concentration (figure 17B) shows a two-step sigmoidal binding-curve suggesting two different binding events of calcium, one low affinity binding site and one high affinity binding site. This is in accordance with the structure of EF5 and EF7 were

EF5 was suggested to have lower affinity to calcium than EF7. The low affinity binding site, might be involved in regulation of PP2A binding to its substrates and to the activity [38] in response to changes in calcium concentration in the cell at certain signaling events. This is in agreement with data published from Magenta et al. [38] who showed that that PR70/PP2A activity increases with Ca^{2+} concentration, and if EF7 is mutated and it cannot bind calcium and activity is lost. To examine whether calcium was required to form the complex between the scaffolding subunit and regulatory subunit, size-exclusion chromatography was used, showing that calcium did not have a significant impact on this interaction.

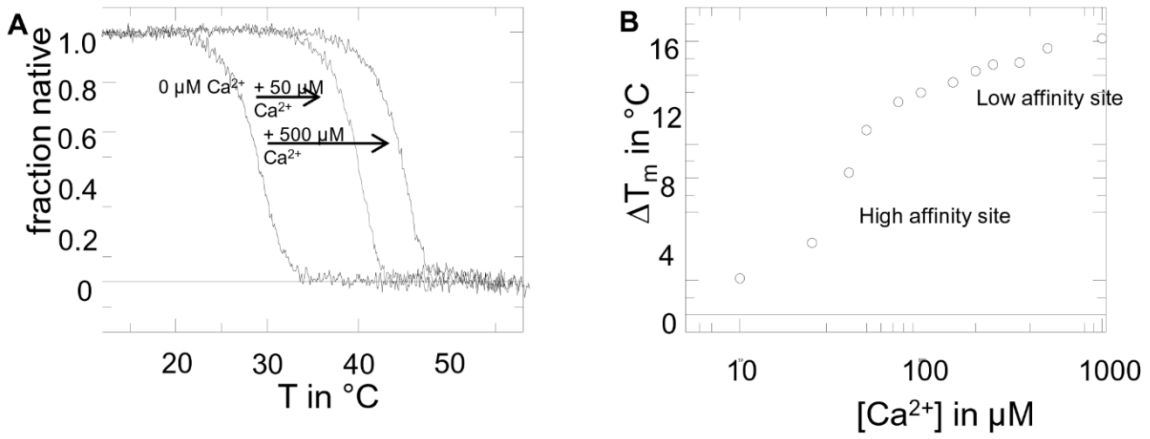

Figure 17. Calcium dependent thermal stability of PR70. A) Representative thermal unfolding curves of the PR70 construct (residues 105–575) are shown in the presence of the different Ca2+ concentrations. B) Ca2+-dependence of thermal unfolding suggests a low and high affinity Ca2+ binding site.

4.5.5 A- and B''-subunit interact with high affinity

A specific feature for the B" regulatory subunit is that they can form a complex with the A subunit in the absence of the catalytic subunit. This does not appear to be the case for other B family members. To study this interaction and map the interaction surfaces of the scaffolding and regulatory subunit, different constructs of A and B" were analyzed with analytical gel filtration (AGF). To be able to detect binding on AGF the affinity has to be quite high, typically in the nM range, so that the complex does not dissociate on the column. AGF showed that HEAT repeat 1-8 of A was enough to bind to B"/PR70, and that constructs with residues 25-575 and 105-575 of B"/PR70 was binding to HEAT repeats 1-8 of A subunit.

To further investigate the binding, isothermal titration calorimetry (ITC) was used to measure the affinity, dissociation constants and change in enthalpy of some constructs (Table 2). In agreement with AGF data, ITC results confirmed that PR70 and the Aα subunit interact with each other with high affinity. The parts of the two subunits that are

important for the binding are residues 105-575 of PR70 and HEAT-repeats 1-8 for Aα, because the interaction is in the nanomolar range, and there is no change in affinity comparing full length constructs with the shorter ones. When truncating PR70 further from the N-terminal by deleting the first 129 residues from the expression construct, the affinity dropped three orders of magnitude. The results indicate that residues 105-129 are important for the binding and belong to the interaction surface. What is also revealed with ITC is that the crystallized construct (PR70 130-483) was also binding to Aα with only very low affinity, which was probably the reason why the complex did not crystallize. To examine whether the C-terminus of PR70 was important for forming a complex with Aα,a construct with only residues 484-575 was designed for the test, no binding could be detected. The conclusion we can draw from the ITC measurements is that residues from 105 to 129 of PR70 definitely belong to the binding surface and that probably some more residues C-terminal of residue 130 also contribute the binding surface.

Table 2. ITC derived binding affinities of B''-constructs to A-subunits

B''-subunit (residues)	A-subunit	N[a]	K_D (μM)	ΔH (kcal/mole)
PR70 (25-575)	A α (HEAT1-15)	0.84 ± 0.02	0.112 ± 0.006	-22.6 ± 0.1
PR70 (105-575)	A α (HEAT1-15)	0.85 ± 0.03	0.045 ± 0.008	-19.7 ± 0.2
PR70 (130-575)	A α (HEAT1-15)		> 100	
PR70 (130-483)	A α (HEAT1-15)		> 100	
PR70 (484-575)	A α (HEAT1-15)		NB	
PR70 (25-575)	A α (HEAT1-8)	0.93 ± 0.02	0.223 ± 0.021	-19.1 ± 0.4
PR70 (105-575)	A α (HEAT1-8)	0.99 ± 0.01	0.158 ± 0.022	-20.3 ± 0.2

4.5.6 And B''- complex with SAXS

Because crystallization attempts of the A-B"–complex was not successful, the structure of the complex was determined at low-resolution with Small Angle X-ray Scattering (SAXS) instead. SAXS studies have the advantage that they are made in solution, so not crystals are needed. Potentially some of the large-scale dynamics of the protein can be monitored, when there are no crystal contacts limiting motions.

With SAXS a calculated mass of a protein that is in accordance with the size of the protein is retrieved. If the calculated mass is much higher than the real mass it gives an indication that the protein is inhomogeneous and potentially partly aggregated. The calculated mass for the A-subunit was 62.5 kDa, which is within the error range from the real molecular weight of 65.3 kDa. The experimental SAXS data of the A-subunit was compared with known structures of A-subunit alone and in complex with the holoenzyme with C, A and PR70(120-478), a homolog of PR72, which was published during the progress of our study was published [41], This gave best fit to the individual A-subunit and worst fit to the A-subunit in the complex, showing that the A-subunit is flexible and can have different conformations depending on what it is scaffolding. The comparison of the different holoenzyme crystal structures (3dw8.pdb[50], 4i5l.pdb[41]) with the obtained SAXS data also showed that the fit was not perfect, indicating that the crystal structures do not reflect all the conformations the A-subunit can form in solution. It was confirmed that the most flexible part of A-subunit is between HEAT repeat 10 and 11, which is the linker between the HEAT repeats binding to regulatory subunit and catalytic subunit. The A-B"complex with SAXS, PR70 (130-575) was subsequently modelled into the structure of holoenzyme with PR70 [41] suggesting that in the complex between the A- and PR70 subunits, the 15 HEAT repeats are still mobile and that the PR70 might be flexible

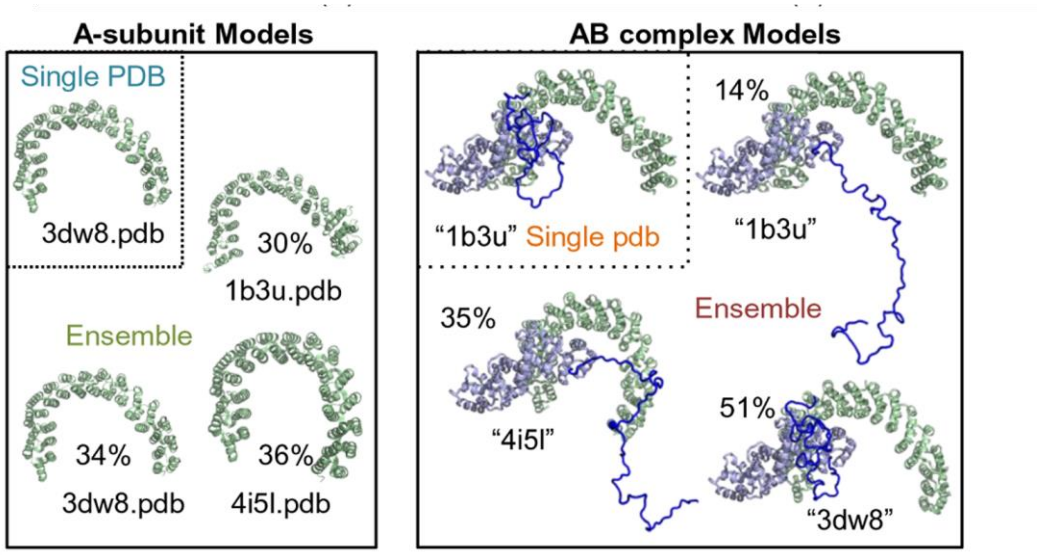

Figure 18. Cartoon depictions of A-subunit models (green) and AB" complex models (green/blue) most consistent with the experimental SAXS data. A-subunit models were taken from representative pdbs. AB" complex models (blue) were based on the published complex structure 4i5l with the A-subunit of 4i5l being replaced with A-subunits from other pdbs. The C-terminus (dark blue) was modeled using BilboMD

5 BIOCHEMICAL PROFILING OF P53

5.1 P53

p53 is a tumor suppressor protein which regulates the cell cycle. It is the most commonly mutated gene identified in human cancers, with 50% of the cancers having a mutation in p53. In response to DNA damage, cell stress and some oncogenic proteins, p53 level rises and induces cell cycle arrest or apoptosis. p53 is acting as a transcription factor [67] to regulate the expression of several different genes. In addition to directly regulate the cell cycle and apoptosis it is implied in cellular process such as metabolic stress sensing, genomic instability, angiogenesis etc. [68].

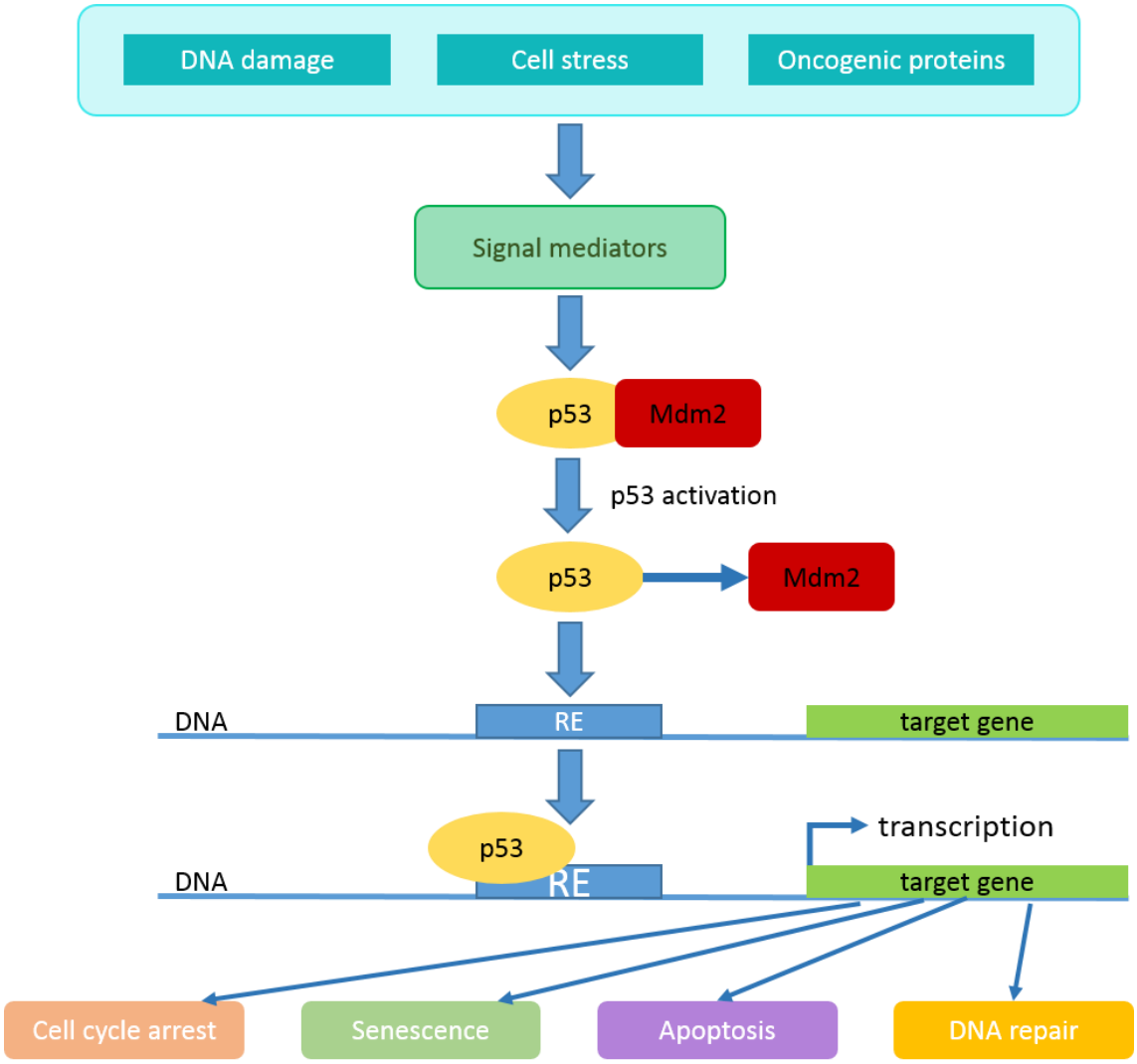

Figure 19. p53 pathway. Due to DNA damage, cell stress, oncogenic proteins etc. signal mediators are activated leading to Mdm2 dissociation from p53, activating p53. p53 binds to a response element (RE) on a target genes that gets transcribed leading to a response like cell cycle arrest, senescence, apoptosis or DNA repair.

5.1.1 DNA-binding

p53 is a DNA-binding protein. The DNA-binding core domain of p53 consists of residues 94-292 that forms an immunoglobulin-like β-sandwich fold and a loop-sheet-helix motif together with two large loops (L2 and L3) [69].The DNA-binding region is stabilized by a Zn ion. p53 binds to a consensus binding site on the DNA of the regulated gene. El-Deiry et al. conducted a precise mapping of the binding sequence which revealed a site with internal symmetry, consisting of two copies of a 10 base-pair motif 5'-RRR**C**(A/T)(T/A)**G**YYY-3' (where R=G and A, and Y=C and T) separated by 0-13 base pairs [70]. The consensus region could be shifted in some p53 target genes; an R gets replaced by a Y and vice versa. Each 10 base-pair motif is called a half-site and the whole binding sequence is called a response element (RE). p53 binds to DNA as a tetramer; two core domains are forming a dimer by binding on each side of a half-site on the DNA double helix, making intermolecular interactions with each other; core-domain to core-domain. The tetramer is then formed as a dimer of dimers. The L3 loop makes interactions with minor groove of DNA backbone via Arg248 in a direct or water-mediated contact and where the Gly245 in the L3 hairpin is essential for the shape of the loop. Residues Ser241 and Arg273 make direct contact with the phosphate backbone of DNA. Arg280 makes a direct contact via two hydrogen bonds to the highly conserved guanine base at position seven in the half-site. Lys120, Ala276 and Cys277 make direct contacts with the major groove of each half-site [71] (figure 20). The carboxyl-terminal of p53 also has ability to bind non-specific DNA in unstressed cells, and helps rapid transactivation of p21 [72].

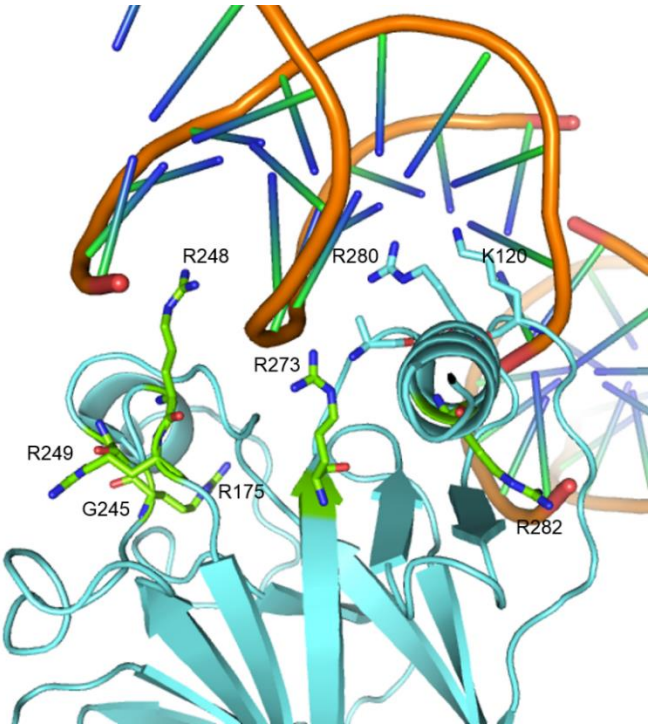

Figure 20. Model p53 core domain binding to DNA. The binding site is rich in basic amino acids. The six most frequently mutated residues are shown in green.

5.1.2 p53 as a tumor suppressor

p53 was first found as a cellular partner of SV40 Large Tumor Antigen, a tumor-virus oncogene [73, 74]. A series of studies after the first finding suggested that p53 was involved in the tumorigenesis and viral replication by interacting with viral proteins from Adenovirus [75], human Papillomavirus [76, 77] and upregulated by T antigen from SV40 [78]. p53 levels are negatively regulated by Mdm2, it is physically blocking the transactivation domain, inhibiting apoptosis and cell cycle arrest [79]. By disrupting Mdm2's ability to bind to p53 by posttranslational modifications like phosphorylation, p53 gets stabilized and activated [80]. After activation, p53 is binding to the response element on the target genes which leads to activation or suppression of the target gene (Figure 20). Some of the target genes that lead to apoptosis are bax, puma, noxa, and apaf-1 [81]. p53 activation and regulation of target genes does not always leads to apoptosis, cell cycle arrest is another outcome. p21, an inhibitor of cyclin-dependent kinases (CDKs) is activated by p53. CDKs are regulating the cell cycle G1-to-S, and G2-to-mitosis steps, hence the inhibition of the CDKs will lead to cell cycle arrest [68]. p53 is not just a tumor suppressor by initiating apoptosis and cell cycle arrest, it is also activating senescence. It is an irreversible cell cycle arrest, where the cells are unable to replicate DNA. One protein regulated by p53 that leads to senescence is plasminogen activator inhibitor 1 (PAI-1) [82] and another is p21 [83]. All three outcomes of the p53 transcriptional program can lead to the suppression of tumors. Apart from apoptosis and senescence, p53 is also capable of contributing to cell survival by targeting proteins that function as apoptosis inhibitors, such as decoy death receptors DcR1 and DcR2 [84]. The effect of p53 activation is decided by many different factors, for example the type of tissue, the nature of the stress signal, and the environment of the cell. The decision whether p53 should act as cell killer or savior could also be dependent on the extent of damage or the duration of stress [85]. Small damages lead to repair of the DNA in contrast to bigger damages that lead to cell death [86].

5.1.3 p53 in cancer

Unlike other tumor suppressor genes that are not expressed in cancer cells, p53 is expressed in cancer cells as the full length protein, but with missense mutations resulting in single amino acid substitutions. Interestingly mutant p53 is often expressed in higher levels than the wild-type [87]. Mutant p53 loses some of its ability to bind to DNA but also tends to gain oncogenic functions, such as interference with wild-type p53-independent apoptosis[88]. The DNA-binding domain of the p53 protein is a "hot spot" for mutation, as the majority of tumor-associated mutations in p53 occur within this region [89]. Six residues; Arg175, Gly245, Arg248, Arg249, Arg273 and Arg282 in the DNA binding domain are frequently mutated. (green in figure 20). The DNA-binding domain is rich in basic residues and mutations on any of the arginine, leading to loss of

DNA binding. Arg273 plays a crucial role in DNA binding by docking p53 to the DNA backbone, mutations of this residue are abundant in human cancer, and leads to dramatic loss of DNA binding affinity [90]. Gly245 is important for the stabilization and folding of p53 as it is located in the L3 loop.

5.2 CETSA STUDIES OF DNA-BINDING TO P53 (PAPER II)

As described above, p53 is binding to DNA at a genes response element (RE). In principle there are many methods for evaluating p53 DNA binding, such as electrophoretic mobility shift assay (EMSA) [91], DNA foot printing [92], ELISA-based techniques with biotinylated REs to capture p53-DNA complexes on streptavidin plates [93], microsphere assay of protein-DNA binding (MAPD) [94] or a multiplex in vitro binding assay quantified with real-time PCR [95]. Other methods used for p53-DNA binding assays are surface plasmon resonance (SPR) and fluorescence anisotropy. All these methods however, require either recombinant p53 or *in vitro* overexpressed p53. We wanted to study p53 in cancer cell lines under a more natural environment without the need of cloning and overexpressing p53. To accomplish this CETSA was used on cancer cell extracts. When oligonucleotides containing REs were added to cell extracts of the cancer cell line (A549) harboring wild-type p53 and one harboring R273H mutant p53 (HT-29), very significant shifts in the T_m for the wild-type p53 were seen but no significant shifts were seen for the mutant p53, suggesting that it is indeed possible to monitor DNA binding events with CETSA in cell extracts (paper II) [22].

5.3 PROFILING OF P53-DNA BINDING (PAPER III)

We wanted to make a proof-of-principle study for the potential to use CETSA to track differences between different mutant and wild-type p53s, and how they bind to REs of different genes. The initial question asked was: does some mutant p53 lose their ability to bind DNA completely or just with decreased affinity and whether some mutants still bind to some gene REs but not others? The profiling study would potentially give a better view how p53 is regulating genes where p53's ability to bind to different REs could shed light on which p53 regulated processes important for the progression of cancer are activated.

5.3.1 Method

The lung carcinoma cell line H1299, a p53 null cell line transfected with wild-type or mutant p53 was kindly provided by David Lane. The cell lines have the same genetic background and the only thing altered is the form of transfected p53. This gives possibility to study only the changes on p53 function so that the DNA binding ability is not affected by other factors. Wild-type and four different p53 mutants, G245S, R248Q, R273C and R273H, were used for the CETSA experiments. Cell extracts from the different cells harboring the p53 variants were incubated with five different

oligonucleotides, of which four contain the response element: p21, PUMA, GADD45 and ConA, and one without the response element as a control: WRNC, as shown in table 3.

Table 3. Oligonucleotide sequences corresponding to REs for genes in the study

Oligo name	Oligonucleotide Sequence
ConA	**CTTTCTGGCCATCATGGGCATGTCCGGGCATGTCCAGCTCTGGCATAGAA**
p21	**CTTTCTGGCCATCATGAACATGTCCCAACATGTTGAGCTCTGGCATAGAA**
PUMA	**CTTTCTGGCCATCATCTGCAAGTCCTGACTTGTCCAGCTCTGGCATAGAA**
GADD45	**CTTTCTGGCCATCATGAACATGTCTAAGCATGCTGAGCTCTGGCATAGAA**
WRNC	**CTTTCTGGCCATCATGAAAGGTGGATTTAGGTGGAAGCTCTGGCATAGAA**

5.3.2 Results

Initially we wanted to investigate the differences in p53 mutants' stability as well as their ability to bind to DNA for different p53 target genes. The results indicate that there were no major differences in the T_m of the different mutants in cell extracts although G245S might be somewhat less stable than the wild-type (figure 21). This later would be in accordance with previous studies, where it was shown that G245S destabilizes the DNA binding core-domain of p53 [96]. The results also demonstrated that not all mutant p53 were interacting to the same extent with different oligonucleotides. For example, PUMA show little stabilization for R273H but binds all other mutants. The results also demonstrated that the mutation in residue 273 to a histidine and cysteine results in somewhat different DNA binding profiles, presumably due to the different size and polarity of these residues.

The DNA binding capabilities of two different cell lines harboring wild-type p53 was studied to see if there were any differences in how wild-type p53 behaves in cancer cells. These were also compared with the different p53 mutants. Since p53 is an important tumor suppressor and is mutated in half of all cancers, it is possible that even wild-type p53 in some cancer cells might change their ability to bind DNA, for example due to other regulating proteins activating or deactivating p53.

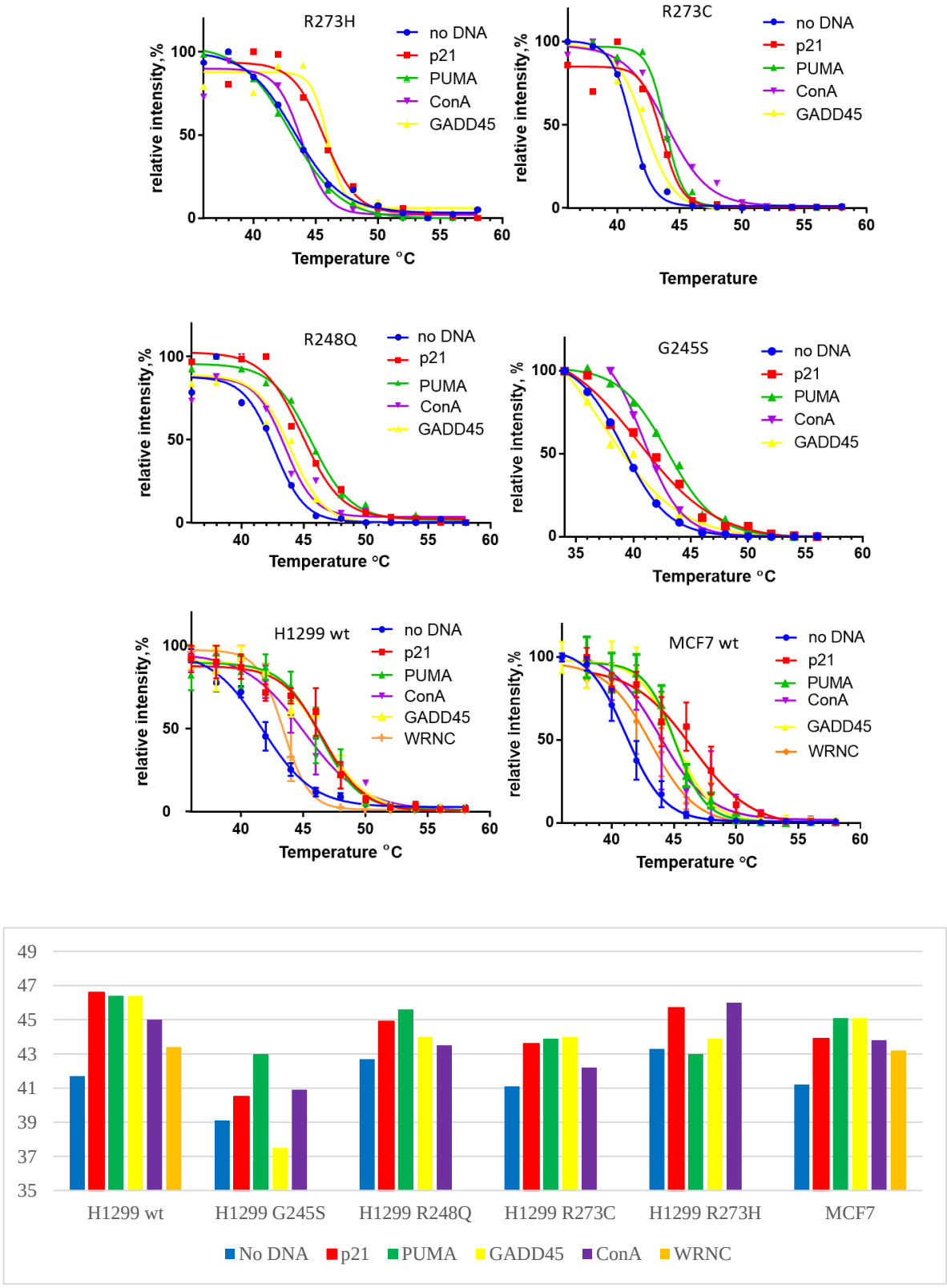

Figure. 21. Stability measurements of mutant and wild-type p53. Heat induced unfolding, detected with western blot. The top shows the melt curves of the mutants and wild-type p53. The diagram is illustrating the T_m.

To test this hypothesis the breast cancer cell line, MCF7, with wild-type p53 and H1299 wild-type was used. The stabilization of p53 in MCF7 appears dampened as compared to H1299 where p53 is transfected into the cell (figure 21). This suggests that there might be an overall difference in DNA binding capabilities between different wild-type p53 harboring cell lines but no significant difference between how they bind to the different oligonucleotides. p53 binds unspecifically to DNA at the C-terminal [97] therefore we also investigated whether it was the unspecific binding of oligonucleotides that was measured in the CETSA experiments. The results showed a slight stabilization using WRNC, an oligonucleotide without an RE, but not in the same extent as with the oligonucleotides containing the RE.

5.4 CETSA FOR GENERATION OF CANCER BIOMARKERS

Understanding how wild-type and mutant p53 binds to its target genes could be helpful in restoring the function of mutant p53 in cancer cells. Several drugs have been developed that can, restore the biological function of p53 [98]), inhibit Mdm2´s ability to bind to p53 [99] or stabilize p53 mutants [100]. Using CETSA coupled with a high throughput detection method like AlphaScreen® it is possible to rapidly screen for drug candidates in compound libraries [23].

In Paper II, CETSA was implemented to study interactions of drugs at the proteome level which has a number of important applications. For example, it is often important that drug candidates are specific for the intended target and does not hit other protein, when such off-target effects can lead to adverse effects. CETSA combined with quantitative mass spectrometry (MS) for thermal profiling of drug targets is in Paper II used for finding off-targets of kinase drugs. When using MS as a detection method for CETSA it is possible to screen a cellular proteome comprising over 7000 proteins for interactions with the drug. Cells are cultured with and without added drug, or cell extract re generated as above, and then heated at different temperatures, for each temperature a sample is taken and the soluble fraction is quantified by mass spectrometry to yield melt curves (Figure 22). From the melt curves, targets that are binding the drug could be identified.

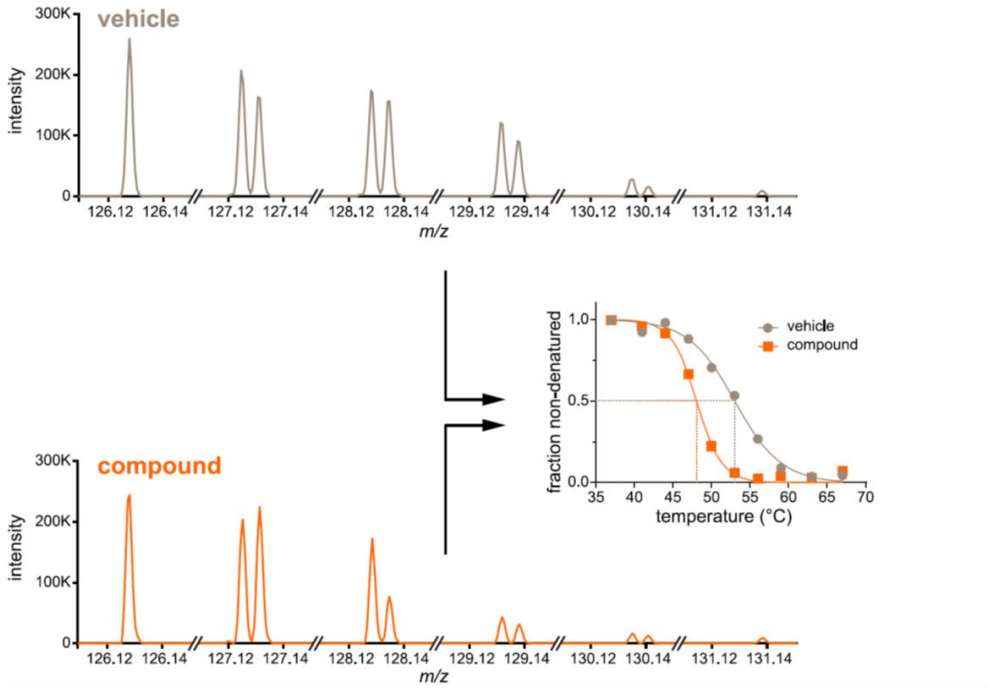

Figure 23. Heated samples are analyzed by LC-MS/MS. The soluble fraction decreases with higher temperatures. The intensities is plotted against temperature to yield melt curves [22].

In this work the heme biosynthesis enzyme ferrochelatase (FECH) was identified to be an off-target of the drug vemurafenib that is developed as a kinase inhibitor targeting BRAF [22]. The work also shows that functional changes in the cell downstream if the drug target BCR-ABL can be directly monitored. This implicates that MS-CETSA could be an important method to identify markers for drug efficacy and toxicity. In the future MS-CETSA could also be used as a method to find cancer biomarkers based on the thermal melting of the proteome that can be used in diagnosis or prognosis.

6 ACKNOWLEDGEMENTS

I would like to take the opportunity to express my gratitude to everyone that has been involved in completing this work and that has been around during my doctoral studies. This has been a real journey and I have learnt so much about both science and life.

First I was to thank Pär Nordlund, my supervisor, for giving me the opportunity to do my PhD studies in your lab. This thesis does definitely not look like I would have expected from the beginning. Thank you for always being an innovator, coming up with new ideas and projects when the first one failed. I have got the great experience in trying everything from method development, to crystallography, biophysical studies and in the end back to method development.

Christian, my co-supervisor. You have really pushed me to become an independent scientist. Changing to the PP2A project saved my thesis. Your methods of teaching is really good. I wish you good look in Hamburg.

Rozbeh, you became my co-supervisor in the end of my studies and taught me all about cell culturing and CETSA. You have also become a friend. A lot of laughter in the fikarum and I will not forget you running around the whole town to find egg free cake for me, and ending up buying Riesen.

Lionel, many many thanks for the cool and awesome cover of my thesis. Good luck with your company ShapeSciFx. I am sure you will get many customers.

Marina, you were my original co-supervisor, teaching me the molecular biology. Thank you for that.

Esben, without you there would not be any structure in this thesis. I think you are one of the most patient people I know.

Fatma and Rachel, the best PhD-student colleagues one could ask for. Our office have had a great environment and you are always so helpful. Rachel, thanks for reading and correcting my thesis. I hope all three of us will stay friends for a long time.

Daniel, thank you for the help with my project. Chit chatting with you about everything and nothing makes my mornings.

Everyone in the lab. Pelle, a great lab manager taking care of all of us. Jens, the number one innovator and entrepreneur, good luck with your company. Raba, ze German business guy, good luck in Munich. Catrine, never stop doing fredagsfys. TC, your cakes a probably tasty. Rosaria, don't forget to rest sometimes. Johan, Christina and Noor.

To everyone at PSF (and former SGC), especially Martin and Thomas for providing the crystallography and biophysics platforms. The ITC machine became my frenemy. Emma, Ida, Eleonora, Helena, and everyone that have left the building.

Herwig's group including: Aja, Torun, Tobias, Filipa, and former members. Thank you for the company during the lunches and fikas in the kitchen.

A big thanks to former members of the lab that have moved on to new adventures: Marie (I know much more about cats now), Malin, Mikaela, Maria, Kerstin (we had so much fun in the office), Madeleine (your journey is awesome), Hanna, Kimmo, Diana, Damian, Daniel, Said, Newsha, Cristine, Cedric.

Thanks to all people I have got to know at MBB and Karolinska for pubs, and other activities.

Tack till släkt och vänner som har funnits där, som inte alltid förstått vad jag sysslar med men ändå hejjat på.

Mamma och pappa. Utan er hade jag aldrig kommit så här långt. Kommer ihåg när jag i slutet av gymnasiet sa att jag skulle bli doktor och ni sa att man inte kan bestämma det bara sådär, man måste hitta ett forskningsprojekt som man brinner för. Sen ändrade jag mig, men till slut hamnade jag här. Ni har alltid pushat mig att vara mitt bästa med era höga förväntningar, tack för det och för att ni tror på mig.

Mina systrar Erica och Maja. Vad skulle jag göra utan er? Systrar gör livet lite roligare och utmanande. Tack för att ni finns. Benjamin och Boris, sötaste systersönerna.

Håkan, mitt livs kärlek. Du är min största supporter och min livscoach. Att du tror på mig mer än jag själv gör ibland är helt ovärderligt. Tack för att du finns vid min sida och lyfter mig när det behövs.

Valeria, min älskade, sötaste, finaste, duktigaste lilla tjej. Du gör livet värt att leva. Jag gör allt för dig. Du ska veta att det finns inga gränser för vad du kan ta dig för att göra i ditt liv. The sky is the limit!

7 REFERENCES

1. Berridge, M.J. *Cell Signalling Biology*. 2014
2. Cooper, G.M., *the Cell: A molecular approach*. 2nd ed2000: Sinauer Associates.
3. Pardee, A.B., *A restriction point for control of normal animal cell proliferation.* Proc Natl Acad Sci U S A, 1974. **71**(4): p. 1286-90.
4. Israels, E.D. and L.G. Israels, *The cell cycle.* Oncologist, 2000. **5**(6): p. 510-3.
5. Miyashita, T. and J.C. Reed, *Tumor suppressor p53 is a direct transcriptional activator of the human bax gene.* Cell, 1995. **80**(2): p. 293-9.
6. Yin, C., et al., *Bax suppresses tumorigenesis and stimulates apoptosis in vivo.* Nature, 1997. **385**(6617): p. 637-40.
7. Cancerfonden, *Cancerfondsrapporten 2014*, 2014.
8. Shi, Y., *Serine/threonine phosphatases: mechanism through structure.* Cell, 2009. **139**(3): p. 468-84.
9. Savitsky, P., et al., *High-throughput production of human proteins for crystallization: the SGC experience.* J Struct Biol, 2010. **172**(1): p. 3-13.
10. Aslanidis, C. and P.J. de Jong, *Ligation-independent cloning of PCR products (LIC-PCR).* Nucleic Acids Res, 1990. **18**(20): p. 6069-74.
11. Lichty, J.J., et al., *Comparison of affinity tags for protein purification.* Protein Expr Purif, 2005. **41**(1): p. 98-105.
12. Porath, J., *Immobilized metal ion affinity chromatography.* Protein Expr Purif, 1992. **3**(4): p. 263-81.
13. Pierce, M.M., C.S. Raman, and B.T. Nall, *Isothermal titration calorimetry of protein-protein interactions.* Methods, 1999. **19**(2): p. 213-21.
14. Phizicky, E.M. and S. Fields, *Protein-protein interactions: methods for detection and analysis.* Microbiol Rev, 1995. **59**(1): p. 94-123.
15. Zopf, D.O., S, *Weak-affinity chromatography.* Nature, 1990. **346**: p. 87–88.
16. Duong-Thi, M.D., et al., *Weak affinity chromatography as a new approach for fragment screening in drug discovery.* Anal Biochem, 2011. **414**(1): p. 138-46.
17. Low, C., et al., *High-throughput analytical gel filtration screening of integral membrane proteins for structural studies.* Biochim Biophys Acta, 2013. **1830**(6): p. 3497-508.
18. Dong, A., et al., *In situ proteolysis for protein crystallization and structure determination.* Nat Methods, 2007. **4**(12): p. 1019-21.
19. *An introduction to circular dichroism spectroscopy*. [cited 2014; Available from: http://www.photophysics.com/tutorials/circular-dichroism-cd-spectroscopy.
20. Vedadi, M., et al., *Biophysical characterization of recombinant proteins: a key to higher structural genomics success.* J Struct Biol, 2010. **172**(1): p. 107-19.

21. Martinez Molina, D., et al., *Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay.* Science, 2013. **341**(6141): p. 84-7.

22. Savitski, M.M., et al., *Proteomics. Tracking cancer drugs in living cells by thermal profiling of the proteome.* Science, 2014. **346**(6205): p. 1255784.

23. Jafari, R., et al., *The cellular thermal shift assay for evaluating drug target interactions in cells.* Nat Protoc, 2014. **9**(9): p. 2100-22.

24. Kendrew, J.C., et al., *A three-dimensional model of the myoglobin molecule obtained by x-ray analysis.* Nature, 1958. **181**(4610): p. 662-6.

25. Abergel, C., *Molecular replacement: tricks and treats.* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 11): p. 2167-73.

26. Svergun, D.I., *Small-angle X-ray and neutron scattering as a tool for structural systems biology.* Biol Chem, 2010. **391**(7): p. 737-43.

27. Svergun, D.I., Koch M.H.J., *Small-angle scattering studies of biological macromolecules in solution.* Rep. Prog. Phys., 2003. **66**: p. 1735–1782.

28. Hunter, T., *Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling.* Cell, 1995. **80**(2): p. 225-36.

29. Wurzenberger, C. and D.W. Gerlich, *Phosphatases: providing safe passage through mitotic exit.* Nat Rev Mol Cell Biol, 2011. **12**(8): p. 469-82.

30. Fischer, E.H. and E.G. Krebs, *Conversion of phosphorylase b to phosphorylase a in muscle extracts.* J Biol Chem, 1955. **216**(1): p. 121-32.

31. Krebs, E.G. and E.H. Fischer, *The phosphorylase b to a converting enzyme of rabbit skeletal muscle.* Biochim Biophys Acta, 1956. **20**(1): p. 150-7.

32. Krebs, E.G., A.B. Kent, and E.H. Fischer, *The muscle phosphorylase b kinase reaction.* J Biol Chem, 1958. **231**(1): p. 73-83.

33. Cori, G.T., Green, A.A., *Crystalline muscle phosphorylase. II. Prosthetic group.* J Biol Chem, 1943. **151**: p. 31-38.

34. Sablina, A.A., et al., *Identification of PP2A complexes and pathways involved in cell transformation.* Cancer Res, 2010. **70**(24): p. 10474-84.

35. Mumby, M.C. and G. Walter, *Protein serine/threonine phosphatases: structure, regulation, and functions in cell growth.* Physiol Rev, 1993. **73**(4): p. 673-99.

36. Gallego, M. and D.M. Virshup, *Protein serine/threonine phosphatases: life, death, and sleeping.* Curr Opin Cell Biol, 2005. **17**(2): p. 197-202.

37. Lechward, K., et al., *Protein phosphatase 2A: variety of forms and diversity of functions.* Acta Biochim Pol, 2001. **48**(4): p. 921-33.

38. Magenta, A., et al., *Protein phosphatase 2A subunit PR70 interacts with pRb and mediates its dephosphorylation.* Mol Cell Biol, 2008. **28**(2): p. 873-82.

39. Kurimchak, A. and X. Grana, *PP2A holoenzymes negatively and positively regulate cell cycle progression by dephosphorylating pocket proteins and multiple CDK substrates.* Gene, 2012. **499**(1): p. 1-7.

40. Cobrinik, D., *Pocket proteins and cell cycle control.* Oncogene, 2005. **24**(17): p. 2796-809.

41. Wlodarchak, N., et al., *Structure of the Ca2+-dependent PP2A heterotrimer and insights into Cdc6 dephosphorylation.* Cell Res, 2013. **23**(7): p. 931-46.

42. Davis, A.J., et al., *Protein phosphatase 2A is targeted to cell division control protein 6 by a calcium-binding regulatory subunit.* J Biol Chem, 2008. **283**(23): p. 16104-14.

43. Cohen, P. and P.T. Cohen, *Protein phosphatases come of age.* J Biol Chem, 1989. **264**(36): p. 21435-8.

44. Wang, S.S., et al., *Alterations of the PPP2R1B gene in human lung and colon cancer.* Science, 1998. **282**(5387): p. 284-7.

45. Calin, G.A., et al., *Low frequency of alterations of the alpha (PPP2R1A) and beta (PPP2R1B) isoforms of the subunit A of the serine-threonine phosphatase 2A in human neoplasms.* Oncogene, 2000. **19**(9): p. 1191-5.

46. Ruediger, R., H.T. Pham, and G. Walter, *Disruption of protein phosphatase 2A subunit interaction in human cancers with mutations in the A alpha subunit gene.* Oncogene, 2001. **20**(1): p. 10-5.

47. Groves, M.R., et al., *The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs.* Cell, 1999. **96**(1): p. 99-110.

48. Xing, Y., *Structure of Protein Phosphatase 2A Core Enzyme bound to tumour-inducing toxins.* Cell, 2006. **127**: p. 341-353.

49. Xu, Y., et al., *Structure of the protein phosphatase 2A holoenzyme.* Cell, 2006. **127**(6): p. 1239-51.

50. Xu, Y., et al., *Structure of a protein phosphatase 2A holoenzyme: insights into B55-mediated Tau dephosphorylation.* Mol Cell, 2008. **31**(6): p. 873-85.

51. Zhou, J., et al., *Characterization of the Aalpha and Abeta subunit isoforms of protein phosphatase 2A: differences in expression, subunit interaction, and evolution.* Biochem J, 2003. **369**(Pt 2): p. 387-98.

52. Janssens, V., S. Longin, and J. Goris, *PP2A holoenzyme assembly: in cauda venenum (the sting is in the tail).* Trends Biochem Sci, 2008. **33**(3): p. 113-21.

53. Virshup, D.M. and S. Shenolikar, *From promiscuity to precision: protein phosphatases get a makeover.* Mol Cell, 2009. **33**(5): p. 537-45.

54. Shi, Y., *Assembly and structure of protein phosphatase 2A.* Sci China C Life Sci, 2009. **52**(2): p. 135-46.

55. Xing, Y., et al., *Structure of protein phosphatase 2A core enzyme bound to tumor-inducing toxins.* Cell, 2006. **127**(2): p. 341-53.

56. Cho, U.S. and W. Xu, *Crystal structure of a protein phosphatase 2A heterotrimeric holoenzyme.* Nature, 2007. **445**(7123): p. 53-7.

57. Neer, E.J., et al., *The ancient regulatory-protein family of WD-repeat proteins.* Nature, 1994. **371**(6495): p. 297-300.

58. Bennecib, M., et al., *Role of protein phosphatase-2A and -1 in the regulation of GSK-3, cdk5 and cdc2 and the phosphorylation of tau in rat forebrain.* FEBS Lett, 2000. **485**(1): p. 87-93.

59. Gong, C.X., et al., *Phosphorylation of microtubule-associated protein tau is regulated by protein phosphatase 2A in mammalian brain. Implications for neurofibrillary degeneration in Alzheimer's disease.* J Biol Chem, 2000. **275**(8): p. 5535-44.

60. Goedert, M. and M.G. Spillantini, *A century of Alzheimer's disease.* Science, 2006. **314**(5800): p. 777-81.

61. Ruediger, R., J. Ruiz, and G. Walter, *Human cancer-associated mutations in the Aalpha subunit of protein phosphatase 2A increase lung cancer incidence in Aalpha knock-in and knockout mice.* Mol Cell Biol, 2011. **31**(18): p. 3832-44.

62. Strynadka, N.C. and M.N. James, *Crystal structures of the helix-loop-helix calcium-binding proteins.* Annu Rev Biochem, 1989. **58**: p. 951-98.

63. Gifford, J.L., M.P. Walsh, and H.J. Vogel, *Structures and metal-ion-binding properties of the Ca2+-binding helix-loop-helix EF-hand motifs.* Biochem J, 2007. **405**(2): p. 199-221.

64. Cook, W.J., et al., *Structure of a sarcoplasmic calcium-binding protein from amphioxus refined at 2.4 A resolution.* J Mol Biol, 1993. **229**(2): p. 461-71.

65. Vijay-Kumar, S. and W.J. Cook, *Structure of a sarcoplasmic calcium-binding protein from Nereis diversicolor refined at 2.0 A resolution.* J Mol Biol, 1992. **224**(2): p. 413-26.

66. Janssens, V., et al., *Identification and functional analysis of two Ca2+-binding EF-hand motifs in the B"/PR72 subunit of protein phosphatase 2A.* J Biol Chem, 2003. **278**(12): p. 10697-706.

67. Riley, T., et al., *Transcriptional control of human p53-regulated genes.* Nat Rev Mol Cell Biol, 2008. **9**(5): p. 402-12.

68. Vogelstein, B., D. Lane, and A.J. Levine, *Surfing the p53 network.* Nature, 2000. **408**(6810): p. 307-10.

69. Cho, Y., et al., *Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations.* Science, 1994. **265**(5170): p. 346-55.

70. el-Deiry, W.S., et al., *Definition of a consensus binding site for p53.* Nat Genet, 1992. **1**(1): p. 45-9.

71. Kitayner, M., et al., *Structural basis of DNA recognition by p53 tetramers.* Mol Cell, 2006. **22**(6): p. 741-53.

72. Liu, Y., et al., *Facilitated search for specific genomic targets by p53 C-terminal basic DNA binding domain.* Cancer Biol Ther, 2004. **3**(11): p. 1102-8.

73. Lane, D.P. and L.V. Crawford, *T antigen is bound to a host protein in SV40-transformed cells.* Nature, 1979. **278**(5701): p. 261-3.

74. Linzer, D.I. and A.J. Levine, *Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells.* Cell, 1979. **17**(1): p. 43-52.

75. Sarnow, P., et al., *Adenovirus E1b-58kd tumor antigen and SV40 large tumor antigen are physically associated with the same 54 kd cellular protein in transformed cells.* Cell, 1982. **28**(2): p. 387-94.

76. Scheffner, M., et al., *The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53.* Cell, 1990. **63**(6): p. 1129-36.

77. Werness, B.A., A.J. Levine, and P.M. Howley, *Association of human papillomavirus types 16 and 18 E6 proteins with p53.* Science, 1990. **248**(4951): p. 76-9.

78. Linzer, D.I., W. Maltzman, and A.J. Levine, *The SV40 A gene product is required for the production of a 54,000 MW cellular tumor antigen.* Virology, 1979. **98**(2): p. 308-18.

79. Lin, J., et al., *Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein.* Genes Dev, 1994. **8**(10): p. 1235-46.

80. Zilfou, J.T. and S.W. Lowe, *Tumor suppressive functions of p53.* Cold Spring Harb Perspect Biol, 2009. **1**(5): p. a001883.

81. Shaw, P., et al., *Induction of apoptosis by wild-type p53 in a human colon tumor-derived cell line.* Proc Natl Acad Sci U S A, 1992. **89**(10): p. 4495-9.

82. Kortlever, R.M., P.J. Higgins, and R. Bernards, *Plasminogen activator inhibitor-1 is a critical downstream target of p53 in the induction of replicative senescence.* Nat Cell Biol, 2006. **8**(8): p. 877-84.

83. Brown, J.P., W. Wei, and J.M. Sedivy, *Bypass of senescence after disruption of p21CIP1/WAF1 gene in normal diploid human fibroblasts.* Science, 1997. **277**(5327): p. 831-4.

84. Janicke, R.U., D. Sohn, and K. Schulze-Osthoff, *The dark side of a tumor suppressor: anti-apoptotic p53.* Cell Death Differ, 2008. **15**(6): p. 959-76.

85. Vousden, K.H. and C. Prives, *Blinded by the Light: The Growing Complexity of p53.* Cell, 2009. **137**(3): p. 413-31.

86. Bensaad, K. and K.H. Vousden, *p53: new roles in metabolism.* Trends Cell Biol, 2007. **17**(6): p. 286-91.

87. Hussain, S.P. and C.C. Harris, *Molecular epidemiology of human cancer: contribution of mutation spectra studies of tumor suppressor genes.* Cancer Res, 1998. **58**(18): p. 4023-37.

88. Sigal, A. and V. Rotter, *Oncogenic mutations of the p53 tumor suppressor: the demons of the guardian of the genome.* Cancer Res, 2000. **60**(24): p. 6788-93.

89. Hainaut, P. and M. Hollstein, *p53 and human cancer: the first ten thousand mutations.* Adv Cancer Res, 2000. **77**: p. 81-137.

90. Bullock, A.N. and A.R. Fersht, *Rescuing the function of mutant p53.* Nat Rev Cancer, 2001. **1**(1): p. 68-76.

91. Fried, M. and D.M. Crothers, *Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis.* Nucleic Acids Res, 1981. **9**(23): p. 6505-25.

92. Galas, D.J. and A. Schmitz, *DNAse footprinting: a simple method for the detection of protein-DNA binding specificity.* Nucleic Acids Res, 1978. **5**(9): p. 3157-70.

93. Jagelska, E., et al., *New ELISA technique for analysis of p53 protein/DNA binding properties.* J Immunol Methods, 2002. **267**(2): p. 227-35.

94. Noureddine, M.A., et al., *Probing the functional impact of sequence variation on p53-DNA interactions using a novel microsphere assay for protein-DNA binding with human cell extracts.* PLoS Genet, 2009. **5**(5): p. e1000462.

95. Goh, W., D. Lane, and F. Ghadessy, *Development of a novel multiplex in vitro binding assay to profile p53-DNA interactions.* Cell Cycle, 2010. **9**(15): p. 3030-8.

96. Ang, H.C., et al., *Effects of common cancer mutations on stability and DNA binding of full-length p53 compared with isolated core domains.* J Biol Chem, 2006. **281**(31): p. 21934-41.

97. Bakalkin, G., et al., *p53 binds single-stranded DNA ends through the C-terminal domain and internal DNA segments via the middle domain.* Nucleic Acids Res, 1995. **23**(3): p. 362-9.

98. Lehmann, S., et al., *Targeting p53 in vivo: a first-in-human study with p53-targeting compound APR-246 in refractory hematologic malignancies and prostate cancer.* J Clin Oncol, 2012. **30**(29): p. 3633-9.

99. Khoo, K.H., C.S. Verma, and D.P. Lane, *Drugging the p53 pathway: understanding the route to clinical efficacy.* Nat Rev Drug Discov, 2014. **13**(3): p. 217-36.

100. Wilcken, R., et al., *Halogen-enriched fragment libraries as leads for drug rescue of mutant p53.* J Am Chem Soc, 2012. **134**(15): p. 6810-8.